GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 01:08:39 ; Search time 15351 Seconds
(without alignments)
11560.501 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

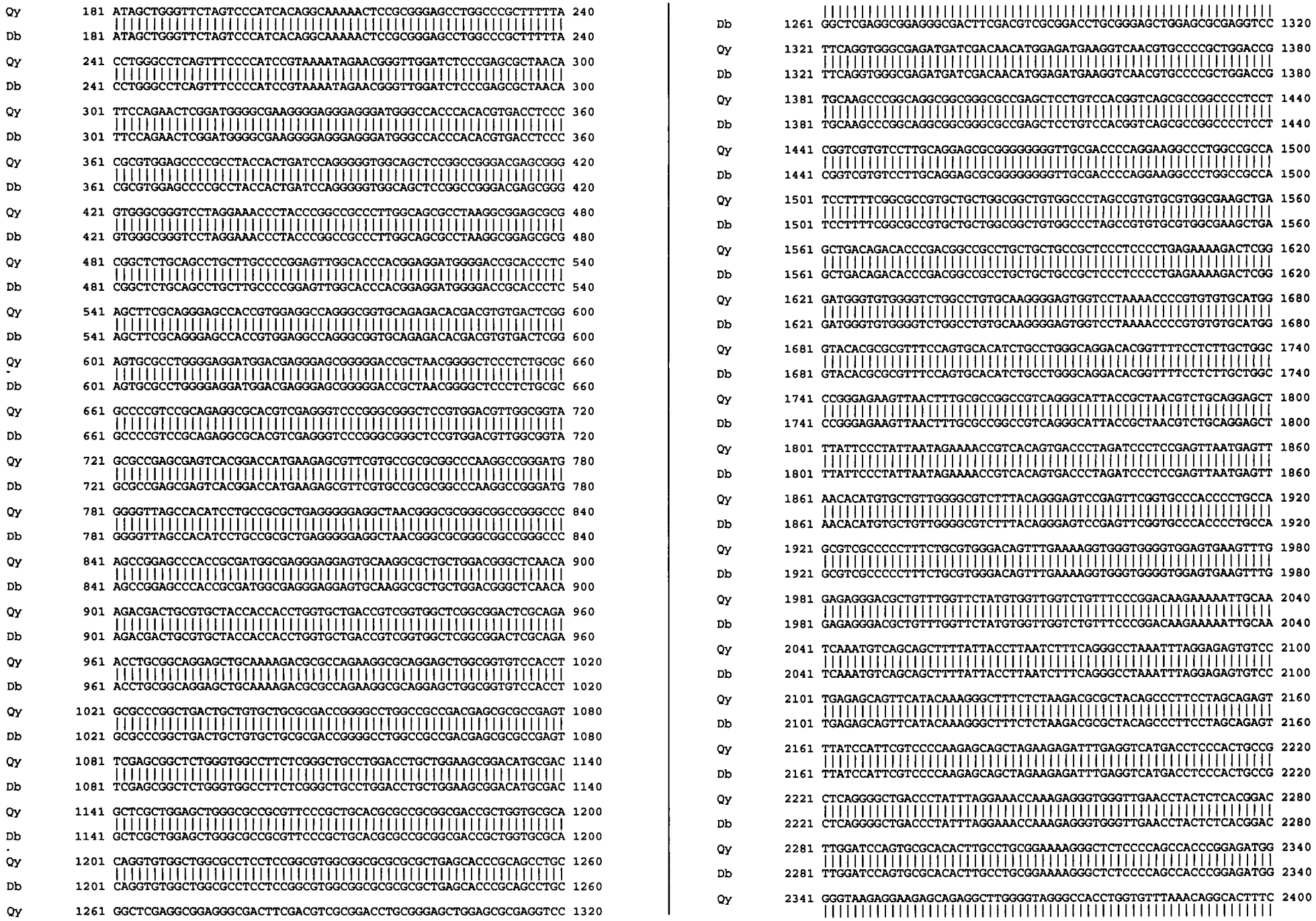
Result No.		Score	% Query Match	Length	DB	ID	Description
1		3122	100.0	3122	6	CQ841455	CQ841455 Sequence
2		3122	100.0	3122	8	AK124499	AK124499 Homo sapi
3	c	2892.8	92.7	191925	8	AC008736	AC008736 Homo sapi
4	c	2892.8	92.7	216441	14	AC027340	AC027340 Homo sapi
5		2860.8	91.6	186115	8	AC008474	AC008474 Homo sapi
6		2308	73.9	2327	6	CQ850499	CQ850499 Sequence
7		2308	73.9	2327	8	AK127646	AK127646 Homo sapi
8		1024.4	32.8	216441	14	AC027340	AC027340 Homo sapi
9		753.8	24.1	211799	9	AC151531	AC151531 Mus muscu
10		753.8	24.1	222531	14	AC166160	AC166160 Mus muscu
11		738	23.6	6560	9	AY229888	AY229888 Mus muscu
12	c	711.8	22.8	255844	14	AC128498	AC128498 Rattus no
13		709.8	22.7	220118	14	AC136661	AC136661 Rattus no
14		706.4	22.6	708	6	CQ748891	CQ748891 Sequence
15		704.8	22.6	708	6	CQ746021	CQ746021 Sequence
16		646	20.7	655	8	HUMZD85E06	AF086461 Homo sapi
17		490.4	15.7	714	9	AF483907	AF483907 Mus muscu
18		480.6	15.4	678	4	AF480475	AF480475 Bos tauru
19		312.6	10.0	4254	5	AY229887	AY229887 Gallus ga
20		312.6	10.0	149625	14	AC150109	AC150109 Gallus ga
21		312.6	10.0	200871	14	AC150036	AC150036 Gallus ga
22	c	263.2	8.4	449	10	G37408	G37408 SHGC-57769
23		235.2	7.5	4238	5	BC084074	BC084074 Xenopus l
24		181.8	5.8	95073	8	AC007272	AC007272 Homo sapi
25	c	181.4	5.8	128461	8	AL390741	AL390741 Human DNA
26	c	181.4	5.8	134443	14	AC015547	AC015547 Homo sapi
27	c	181.2	5.8	63437	8	HSJ964N17	AL050311 Human DNA
28		181.2	5.8	160142	8	AC093684	AC093684 Homo sapi
29	c	181.2	5.8	165167	14	AC010836	AC010836 Homo sapi
30	c	181.2	5.8	211550	8	AC073912	AC073912 Homo sapi
31	c	180.4	5.8	161867	8	AC090513	AC090513 Homo sapi
32	c	180.4	5.8	185979	8	AC091296	AC091296 Pan trogl
33	c	180.4	5.8	245182	14	AC145940	AC145940 Pan trogl
34	c	180.2	5.8	159764	8	AC093866	AC093866 Homo sapi
35	c	180.2	5.8	185371	8	HS625H18	AL022726 Human DNA
36	c	180	5.8	60368	8	CR383701	CR383701 Human DNA
37	c	180	5.8	153087	14	AC017072	AC017072 Homo sapi
38		179.8	5.8	171631	14	AC069291	AC069291 Homo sapi
39	c	179.8	5.8	171788	8	AC073089	AC073089 Homo sapi
40		179.8	5.8	184927	14	AL683875	AL683875 Homo sapi
41		179.8	5.8	189825	8	AL683807	AL683807 Human DNA
42	c	179.8	5.8	193360	14	AC021977	AC021977 Homo sapi
43	c	179.8	5.8	198764	8	AC103809	AC103809 Homo sapi
44		179.6	5.8	102370	8	HSJ976O13	AL117354 Human DNA
45		179.6	5.8	110019	8	AC068860	AC068860 Homo sapi

ALIGNMENTS

RESULT 1
CQ841455
LOCUS CQ841455 3122 bp DNA linear PAT 02-AUG-2004
DEFINITION Sequence 102 from Patent EP1440981.
ACCESSION CQ841455
VERSION CQ841455.1 GI:50893242
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.
TITLE Full-length human cdna
JOURNAL Patent: EP 1440981-A 102 28-JUL-2004;
Research Association for Biotechnology (JP)
FEATURES Location/Qualifiers
source 1. .3122
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 3122; DB 6; Length 3122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
```


Qy 181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
Db 181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
Qy 241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
Db 241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
Qy 301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
Db 301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
Qy 361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
Db 361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
Qy 421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
Db 421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
Qy 481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
Db 481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
Qy 541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
Db 541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
Qy 601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
Db 601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
Qy 661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
Db 661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
Qy 721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
Db 721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
Qy 781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
Db 781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
Qy 841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
Db 841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
Qy 901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
Db 901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
Qy 961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Db 961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Qy 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Db 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Qy 1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Db 1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Qy 1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Db 1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Qy 1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Db 1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Qy 1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Db 1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Qy 1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Db 1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Qy 1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Db 1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Qy 1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Db 1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Qy 1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Db 1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Qy 1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Db 1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Qy 1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Db 1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Qy 1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Db 1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Qy 1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Db 1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Qy 1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Db 1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Qy 1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Db 1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Qy 1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Db 1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Qy 1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Db 1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Qy 2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Db 2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Qy 2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Db 2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Qy 2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Db 2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Qy 2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Db 2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Qy 2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Db 2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Qy 2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

```
Db       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880

Qy       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940

Qy       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000

Qy       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060

Qy       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120

Qy       3121 TC 3122
              ||
Db       3121 TC 3122
```

```
RESULT 2
AK124499
LOCUS       AK124499               3122 bp    mRNA    linear   PRI 09-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ42508 fis, clone BRACE2043665.
ACCESSION   AK124499
VERSION     AK124499.1  GI:34530295
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3122)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .3122
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRACE2043665"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE2"
                     /note="cloning vector: pME18SFL3"
     CDS             858. .1565
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC85865.1"
                     /db_xref="GI:34530296"
                     /translation="MAREECKALLDGLNKTTACYHHLVLTVGGSADSQNLRQELQKTR
                     QKAQELAVSTCARLTAVLRDRGLAADERAEFERLWVAFSGCLDLLEADMRRSLELGAA
                     FPLHAPRRPLVRTGVAGASSGVAARALSTRSLRLEAEGDFDVADLRELEREVLQVGEM
                     IDNMEMKVNVPRWTVQARQAAGAELLSTVSAGPSSVVSLQERGGGCDPRKALAAILFG
                     AVLLAAVALAVCVAKLS"
ORIGIN

  Query Match           100.0%;  Score 3122;  DB 8;  Length 3122;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3122;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440

Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920

Qy       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
```

```
Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 2880

Qy       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940

Qy       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000

Qy       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060

Qy       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120

Qy       3121 TC 3122
              ||
Db       3121 TC 3122
```

```
RESULT 3
AC008736/c
LOCUS       AC008736               191925 bp    DNA     linear   PRI 27-SEP-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.
ACCESSION   AC008736
VERSION     AC008736.6  GI:10312244
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 27, 2000 this sequence version replaced gi:8575905.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
            STS Content:
            SHGC-57769 G37408.
FEATURES             Location/Qualifiers
     source          1. .191925
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2538C1"
ORIGIN
```

Query Match 92.7%; Score 2892.8; DB 8; Length 191925;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129460 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 129401

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129400 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 129341

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129340 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 129281

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129280 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 129221

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129220 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 129161

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129160 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 129101

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129100 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 129041

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129040 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 128981

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128980 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 128921

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128920 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 128861

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128860 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 128801

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128800 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 128741

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128740 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 128681

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128680 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 128621

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128620 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 128561

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
```

```
Db    128560 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 128501

Qy       961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Db    128500 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 128441

Qy      1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Db    128440 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 128381

Qy      1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Db    128380 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 128321

Qy      1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Db    128320 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 128261

Qy      1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Db    128260 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 128201

Qy      1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Db    128200 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 128141

Qy      1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Db    128140 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 128081

Qy      1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Db    128080 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 128021

Qy      1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Db    128020 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 127961

Qy      1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Db    127960 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 127901

Qy      1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Db    127900 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 127841

Qy      1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Db    127840 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 127781

Qy      1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Db    127780 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 127721

Qy      1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Db    127720 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 127661

Qy      1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Db    127660 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 127601

Qy      1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Db    127600 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 127541

Qy      1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Db    127540 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 127481

Qy      1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
```

```
Db    127480 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 127421

Qy      2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Db    127420 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 127361

Qy      2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Db    127360 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 127301

Qy      2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Db    127300 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 127241

Qy      2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Db    127240 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 127181

Qy      2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Db    127180 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 127121

Qy      2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
Db    127120 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 127061

Qy      2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
Db    127060 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 127001

Qy      2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
Db    127000 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 126941

Qy      2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
Db    126940 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 126881

Qy      2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
Db    126880 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 126821

Qy      2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
Db    126820 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 126761

Qy      2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
Db    126760 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 126701

Qy      2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
Db    126700 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 126641

Qy      2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 2880
Db    126640 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 126581

Qy      2881 TCTTTTTTGAGGCAGG 2896
Db    126580 TCTTTTTTGAGACAGG 126565
```

```
RESULT 4
AC027340/c
LOCUS       AC027340               216441 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT
            SEQUENCE, 52 unordered pieces.
ACCESSION   AC027340
VERSION     AC027340.2  GI:9211228
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 15, 2000 this sequence version replaced gi:7341654.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 726304, BC691328
            Center clone name: CITB-H1_2257C19
            -------
            Summary Statistics
            Consensus quality: 164680 bases at least Q40
            Consensus quality: 187929 bases at least Q30
            Consensus quality: 195340 bases at least Q20
            Estimated insert size: 159060; agarose-fp estimation
            Estimated insert size: 211341; sum-of-contigs estimation
            Quality coverage:  8.59 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.46 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1041: contig of 1041 bp in length
            *     1042     1141: gap of unknown length
            *     1142     2370: contig of 1229 bp in length
            *     2371     2470: gap of unknown length
            *     2471     3552: contig of 1082 bp in length
            *     3553     3652: gap of unknown length
            *     3653     4683: contig of 1031 bp in length
            *     4684     4783: gap of unknown length
            *     4784     6178: contig of 1395 bp in length
            *     6179     6278: gap of unknown length
            *     6279     7442: contig of 1164 bp in length
            *     7443     7542: gap of unknown length
            *     7543     8661: contig of 1119 bp in length
            *     8662     8761: gap of unknown length
            *     8762    10030: contig of 1269 bp in length
            *    10031    10130: gap of unknown length
            *    10131    11491: contig of 1361 bp in length
            *    11492    11591: gap of unknown length
            *    11592    12967: contig of 1376 bp in length
            *    12968    13067: gap of unknown length
            *    13068    14173: contig of 1106 bp in length
            *    14174    14273: gap of unknown length
            *    14274    15840: contig of 1567 bp in length
            *    15841    15940: gap of unknown length
            *    15941    17242: contig of 1302 bp in length
            *    17243    17342: gap of unknown length
            *    17343    18449: contig of 1107 bp in length
            *    18450    18549: gap of unknown length
            *    18550    19683: contig of 1134 bp in length
            *    19684    19783: gap of unknown length
            *    19784    21247: contig of 1464 bp in length
            *    21248    21347: gap of unknown length
            *    21348    22952: contig of 1605 bp in length
            *    22953    23052: gap of unknown length
            *    23053    24452: contig of 1400 bp in length
            *    24453    24552: gap of unknown length
            *    24553    25825: contig of 1273 bp in length
            *    25826    25925: gap of unknown length
            *    25926    27109: contig of 1184 bp in length
            *    27110    27209: gap of unknown length
            *    27210    28292: contig of 1083 bp in length
            *    28293    28392: gap of unknown length
            *    28393    29744: contig of 1352 bp in length
            *    29745    29844: gap of unknown length
            *    29845    31042: contig of 1198 bp in length
            *    31043    31142: gap of unknown length
            *    31143    32629: contig of 1487 bp in length
            *    32630    32729: gap of unknown length
            *    32730    33898: contig of 1169 bp in length
            *    33899    33998: gap of unknown length
            *    33999    35704: contig of 1706 bp in length
            *    35705    35804: gap of unknown length
            *    35805    37471: contig of 1667 bp in length
            *    37472    37571: gap of unknown length
            *    37572    39285: contig of 1714 bp in length
            *    39286    39385: gap of unknown length
            *    39386    40577: contig of 1192 bp in length
            *    40578    40677: gap of unknown length
            *    40678    41930: contig of 1253 bp in length
            *    41931    42030: gap of unknown length
            *    42031    43743: contig of 1713 bp in length
            *    43744    43843: gap of unknown length
            *    43844    46054: contig of 2211 bp in length
            *    46055    46154: gap of unknown length
            *    46155    49379: contig of 3225 bp in length
            *    49380    49479: gap of unknown length
            *    49480    53371: contig of 3892 bp in length
            *    53372    53471: gap of unknown length
            *    53472    55756: contig of 2285 bp in length
            *    55757    55856: gap of unknown length
            *    55857    58019: contig of 2163 bp in length
            *    58020    58119: gap of unknown length
            *    58120    60327: contig of 2208 bp in length
            *    60328    60427: gap of unknown length
            *    60428    64223: contig of 3796 bp in length
            *    64224    64323: gap of unknown length
            *    64324    67218: contig of 2895 bp in length
            *    67219    67318: gap of unknown length
            *    67319    71797: contig of 4479 bp in length
            *    71798    71897: gap of unknown length
            *    71898    76856: contig of 4959 bp in length
            *    76857    76956: gap of unknown length
            *    76957    82060: contig of 5104 bp in length
            *    82061    82160: gap of unknown length
            *    82161    87557: contig of 5397 bp in length
            *    87558    87657: gap of unknown length
            *    87658    94226: contig of 6569 bp in length
            *    94227    94326: gap of unknown length
            *    94327   101878: contig of 7552 bp in length
            *   101879   101978: gap of unknown length
            *   101979   110629: contig of 8651 bp in length
            *   110630   110729: gap of unknown length
            *   110730   119083: contig of 8354 bp in length
            *   119084   119183: gap of unknown length
            *   119184   128919: contig of 9736 bp in length
            *   128920   129019: gap of unknown length
            *   129020   142791: contig of 13772 bp in length
            *   142792   142891: gap of unknown length
            *   142892   156079: contig of 13188 bp in length
            *   156080   156179: gap of unknown length
            *   156180   183092: contig of 26913 bp in length
            *   183093   183192: gap of unknown length
            *   183193   216441: contig of 33249 bp in length.
FEATURES             Location/Qualifiers
     source          1. .216441
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="19"
                     /clone="CTD-2257C19"
                     /clone_lib="CalTech human BAC library D"
     gap             1042. .1141
                     /estimated_length=unknown
     gap             2371. .2470
                     /estimated_length=unknown
     gap             3553. .3652
                     /estimated_length=unknown
     gap             4684. .4783
                     /estimated_length=unknown
     gap             6179. .6278
                     /estimated_length=unknown
     gap             7443. .7542
                     /estimated_length=unknown
     gap             8662. .8761
                     /estimated_length=unknown
     gap             10031. .10130
                     /estimated_length=unknown
     gap             11492. .11591
                     /estimated_length=unknown
     gap             12968. .13067
                     /estimated_length=unknown
     gap             14174. .14273
                     /estimated_length=unknown
     gap             15841. .15940
                     /estimated_length=unknown
     gap             17243. .17342
                     /estimated_length=unknown
     gap             18450. .18549
                     /estimated_length=unknown
     gap             19684. .19783
                     /estimated_length=unknown
     gap             21248. .21347
                     /estimated_length=unknown
     gap             22953. .23052
                     /estimated_length=unknown
     gap             24453. .24552
                     /estimated_length=unknown
     gap             25826. .25925
                     /estimated_length=unknown
     gap             27110. .27209
                     /estimated_length=unknown
     gap             28293. .28392
                     /estimated_length=unknown
     gap             29745. .29844

  Query Match             92.7%;  Score 2892.8;  DB 14;  Length 216441;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2894;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117034 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 116975

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116974 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 116915

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116914 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 116855

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116854 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 116795

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116794 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 116735

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     116734 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 116675

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116674 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 116615

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116614 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 116555

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116554 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 116495

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116494 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 116435

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116434 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 116375

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116374 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 116315

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116314 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 116255

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116254 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 116195

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116194 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 116135

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116134 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 116075

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116074 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 116015

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116014 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 115955

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115954 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 115895

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115894 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 115835

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115834 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 115775

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115774 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 115715

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115714 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 115655

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115654 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 115595
```

```
Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115594 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 115535

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115534 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 115475

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115474 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 115415

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115414 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 115355

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115354 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 115295

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115294 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 115235

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115234 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 115175

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115174 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 115115

Qy       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115114 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 115055

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115054 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 114995

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114994 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 114935

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114934 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 114875

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114874 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 114815

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114814 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 114755

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114754 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 114695

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114694 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 114635

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114634 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 114575

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114574 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 114515

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114514 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 114455

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114454 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 114395

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114394 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 114335

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114334 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 114275

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114274 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 114215

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114214 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 114155

Qy       2881 TCTTTTTTGAGGCAGG 2896
             |||||||||||| ||||
Db     114154 TCTTTTTTGAGACAGG 114139
```

```
RESULT 5
AC008474
LOCUS       AC008474              186115 bp    DNA     linear   PRI 13-JUL-2002
DEFINITION  Homo sapiens chromosome 19 clone CTC-379B2, complete sequence.
ACCESSION   AC008474
VERSION     AC008474.9  GI:21743729
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 13, 2002 this sequence version replaced gi:16197759.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.4.
FEATURES             Location/Qualifiers
```

```
     source          1. .186115
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTC-379B2"
ORIGIN

  Query Match             91.6%;  Score 2860.8;  DB 8;  Length 186115;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2887;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180657 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 180716

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180717 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 180776

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180777 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180836

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180837 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 180896

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180897 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 180956

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180957 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 181016

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181017 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 181076

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181077 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 181136

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181137 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 181196

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181197 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 181256

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181257 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 181316

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181317 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 181376

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181377 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 181436

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181437 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 181496

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181497 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 181556

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181557 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 181616

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181617 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 181676

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181677 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 181736

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181737 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 181796

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              || ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     181797 GCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 181856

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181857 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 181916

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181917 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 181976

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181977 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 182036

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182037 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 182096

Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182097 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 182156

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182157 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 182216

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182217 GCTGACGGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 182276

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182277 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 182336

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     182337 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGC 182395

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182396 CCGGGAGGAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 182455

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182456 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 182515

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182516 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 182575

Qy       1921 GCGTCGCCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182576 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 182635

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     182636 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 182695

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182696 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 182755

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182756 CGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 182815

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182816 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 182875

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182876 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 182935

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182936 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 182995

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182996 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 183055

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183056 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 183115

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183116 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 183175

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183176 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 183235

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183236 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 183295

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183296 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATACTCTGAAA 183355

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183356 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 183415

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183416 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 183475

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db     183476 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTT-CCCCCCCCCCCCACTTTTT 183534

Qy       2881 TCTTTTTTGAGGCAGG 2896
              |||||||||| ||||
Db     183535 TCTTTTTTGAGACAGG 183550


RESULT 6
CQ850499
LOCUS       CQ850499               2327 bp    DNA     linear   PAT 23-AUG-2004
DEFINITION  Sequence 968 from Patent EP1447413.
ACCESSION   CQ850499
VERSION     CQ850499.1  GI:51508711
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,
            Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
  TITLE     Full-length human cDNA
  JOURNAL   Patent: EP 1447413-A 968 18-AUG-2004;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .2327
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             73.9%;  Score 2308;  DB 6;  Length 2327;
  Best Local Similarity   99.7%;  Pred. No. 4.4e-290;
  Matches 2322;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 60

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 120

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 180

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 240

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 300

Qy        744 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 360

Qy        804 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 420

Qy        864 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 480

Qy        924 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 540

Qy        984 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 600

Qy       1044 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 660

Qy       1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        661 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGCGCCGCG 720

Qy       1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        721 TTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 780

Qy       1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        781 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 840

Qy       1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 900

Qy       1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 960

Qy       1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1020

Qy       1464 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1080

Qy       1524 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGC 1583
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1081 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACGGACACCCGACGGCCGC 1140

Qy       1584 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1200

Qy       1644 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1260

Qy       1704 ATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCC 1763
             |||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Db       1261 ATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGCCCGGGAGGAGTTAACTTTGCGCC 1319

Qy       1764 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1379

Qy       1824 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1439

Qy       1884 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1499

Qy       1944 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 2003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 1559

Qy       2004 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 2063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 1619

Qy       2064 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGC 2123
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1620 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGC 1679

Qy       2124 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 2183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 1739

Qy       2184 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 2243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 1799

Qy       2244 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 2303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 1859

Qy       2304 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 2363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 1919

Qy       2364 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 2423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 1979

Qy       2424 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2039

Qy       2484 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2099

Qy       2544 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGGTAATGTGGCATTACTGGCCCA 2603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGGTAATGTGGCATTACTGGCCCA 2159

Qy       2604 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2219

Qy       2664 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2279

Qy       2724 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2771
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2327
```

```
RESULT 7
AK127646
LOCUS       AK127646               2327 bp    mRNA    linear   PRI 19-FEB-2004
DEFINITION  Homo sapiens cDNA FLJ45744 fis, clone KIDNE2017153.
ACCESSION   AK127646
VERSION     AK127646.1  GI:34534649
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
            Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and
            Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2327)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: Reverse Proteomics Research Institute, HRI and
            RAB.
FEATURES             Location/Qualifiers
     source          1. .2327
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

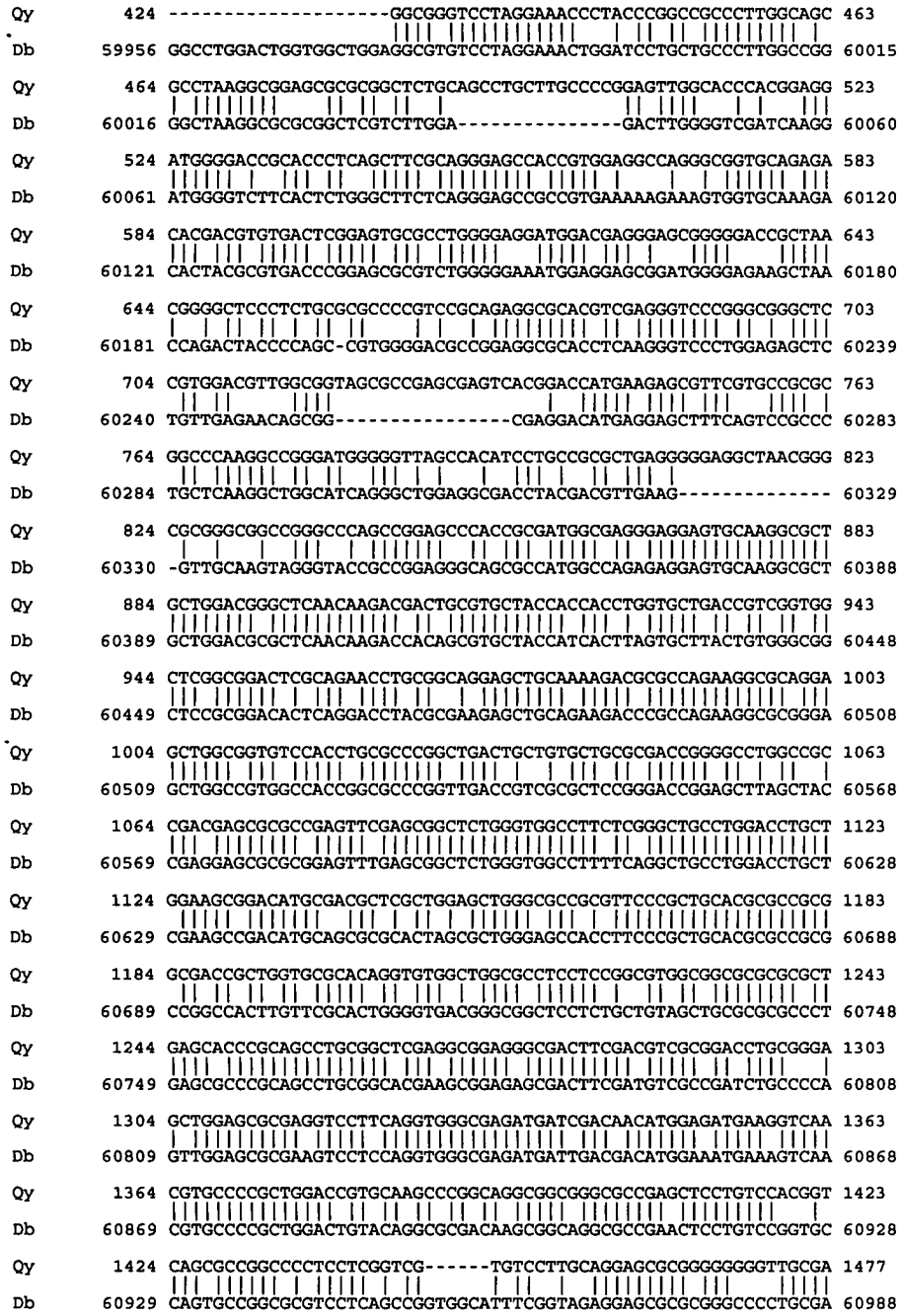
```
                     /clone="KIDNE2017153"
                     /tissue_type="kidney"
                     /clone_lib="KIDNE2"
                     /note="cloning vector: pME18SFL3"
     CDS             415. .1122
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC87072.1"
                     /db_xref="GI:34534650"
                     /translation="MAREECKALLDGLNKTTACYHHLVLTVGGSADSQNLRQELQKTR
                     QKAQELAVSTCARLTAVLRDRGLAADERAEFERLWVAFSGCLDLLEADMRRALELGAA
                     FPLHAPRRPLVRTGVAGASSGVAARALSTRSLRLEAEGDFDVADLRELEREVLQVGEM
                     IDNMEMKVNVPRWTVQARQAAGAELLSTVSAGPSSVVSLQERGGGCDPRKALAAILFG
                     AVLLAAVALAVCVAKLS"
ORIGIN

  Query Match          73.9%;  Score 2308;  DB 8;  Length 2327;
  Best Local Similarity   99.7%;  Pred. No. 4.4e-290;
  Matches 2322;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
Db         1 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 60

Qy       504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
Db        61 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 120

Qy       564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
Db       121 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 180

Qy       624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
Db       181 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 240

Qy       684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
Db       241 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 300

Qy       744 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 803
Db       301 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 360

Qy       804 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 863
Db       361 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 420

Qy       864 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 923
Db       421 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 480

Qy       924 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 983
Db       481 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 540

Qy       984 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 1043
Db       541 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 600

Qy      1044 CGCGACCGGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
Db       601 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 660

Qy      1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163
Db       661 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGCGCCGCG 720

Qy      1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223
Db       721 TTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 780

Qy      1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283
Db       781 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 840

Qy      1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343
Db       841 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 900

Qy      1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403
Db       901 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 960

Qy      1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1463
Db       961 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1020

Qy      1464 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1523
Db      1021 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1080

Qy      1524 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGC 1583
Db      1081 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACGGACACCCGACGGCCGC 1140

Qy      1584 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1643
Db      1141 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1200

Qy      1644 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1703
Db      1201 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1260

Qy      1704 ATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCC 1763
Db      1261 ATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGCCCGGGAGGAGTTAACTTTGCGCC 1319

Qy      1764 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1823
Db      1320 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1379

Qy      1824 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1883
Db      1380 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1439

Qy      1884 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1943
Db      1440 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1499

Qy      1944 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 2003
Db      1500 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 1559

Qy      2004 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 2063
Db      1560 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 1619

Qy      2064 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGC 2123
Db      1620 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGC 1679

Qy      2124 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 2183
Db      1680 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 1739

Qy      2184 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 2243
Db      1740 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 1799

Qy      2244 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 2303
Db      1800 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 1859

Qy      2304 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 2363
```

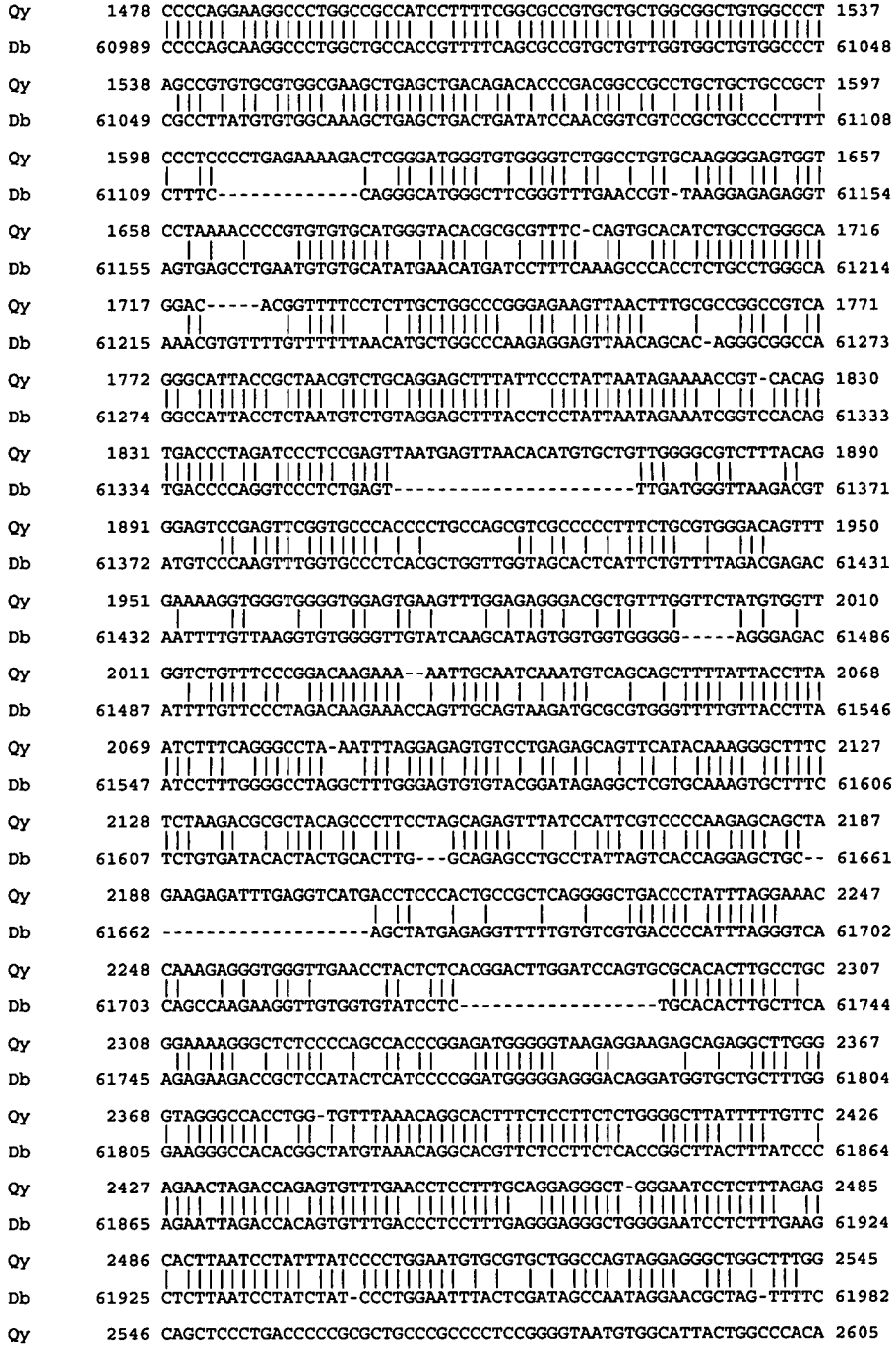
```
Db       1860 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 1919

Qy       2364 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 2423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 1979

Qy       2424 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2039

Qy       2484 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2099

Qy       2544 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2159

Qy       2604 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2219

Qy       2664 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2279

Qy       2724 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2771
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2327


RESULT 8
AC027340
LOCUS       AC027340               216441 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT
            SEQUENCE, 52 unordered pieces.
ACCESSION   AC027340
VERSION     AC027340.2  GI:9211228
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 15, 2000 this sequence version replaced gi:7341654.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 726304, BC691328
            Center clone name: CITB-H1_2257C19
            -------
            Summary Statistics
            Consensus quality: 164680 bases at least Q40
            Consensus quality: 187929 bases at least Q30
            Consensus quality: 195340 bases at least Q20
            Estimated insert size: 159060; agarose-fp estimation
            Estimated insert size: 211341; sum-of-contigs estimation
            Quality coverage:  8.59 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.46 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1041: contig of 1041 bp in length
            *     1042     1141: gap of unknown length
            *     1142     2370: contig of 1229 bp in length
            *     2371     2470: gap of unknown length
            *     2471     3552: contig of 1082 bp in length
            *     3553     3652: gap of unknown length
            *     3653     4683: contig of 1031 bp in length
            *     4684     4783: gap of unknown length
            *     4784     6178: contig of 1395 bp in length
            *     6179     6278: gap of unknown length
            *     6279     7442: contig of 1164 bp in length
            *     7443     7542: gap of unknown length
            *     7543     8661: contig of 1119 bp in length
            *     8662     8761: gap of unknown length
            *     8762    10030: contig of 1269 bp in length
            *    10031    10130: gap of unknown length
            *    10131    11491: contig of 1361 bp in length
            *    11492    11591: gap of unknown length
            *    11592    12967: contig of 1376 bp in length
            *    12968    13067: gap of unknown length
            *    13068    14173: contig of 1106 bp in length
            *    14174    14273: gap of unknown length
            *    14274    15840: contig of 1567 bp in length
            *    15841    15940: gap of unknown length
            *    15941    17242: contig of 1302 bp in length
            *    17243    17342: gap of unknown length
            *    17343    18449: contig of 1107 bp in length
            *    18450    18549: gap of unknown length
            *    18550    19683: contig of 1134 bp in length
            *    19684    19783: gap of unknown length
            *    19784    21247: contig of 1464 bp in length
            *    21248    21347: gap of unknown length
            *    21348    22952: contig of 1605 bp in length
            *    22953    23052: gap of unknown length
            *    23053    24452: contig of 1400 bp in length
            *    24453    24552: gap of unknown length
            *    24553    25825: contig of 1273 bp in length
            *    25826    25925: gap of unknown length
            *    25926    27109: contig of 1184 bp in length
            *    27110    27209: gap of unknown length
            *    27210    28292: contig of 1083 bp in length
            *    28293    28392: gap of unknown length
            *    28393    29744: contig of 1352 bp in length
            *    29745    29844: gap of unknown length
            *    29845    31042: contig of 1198 bp in length
            *    31043    31142: gap of unknown length
            *    31143    32629: contig of 1487 bp in length
            *    32630    32729: gap of unknown length
            *    32730    33898: contig of 1169 bp in length
            *    33899    33998: gap of unknown length
            *    33999    35704: contig of 1706 bp in length
            *    35705    35804: gap of unknown length
            *    35805    37471: contig of 1667 bp in length
            *    37472    37571: gap of unknown length
            *    37572    39285: contig of 1714 bp in length
            *    39286    39385: gap of unknown length
            *    39386    40577: contig of 1192 bp in length
            *    40578    40677: gap of unknown length
            *    40678    41930: contig of 1253 bp in length
            *    41931    42030: gap of unknown length
            *    42031    43743: contig of 1713 bp in length
            *    43744    43843: gap of unknown length
            *    43844    46054: contig of 2211 bp in length
            *    46055    46154: gap of unknown length
            *    46155    49379: contig of 3225 bp in length
            *    49380    49479: gap of unknown length
            *    49480    53371: contig of 3892 bp in length
```

```
          *    53372    53471: gap of unknown length
          *    53472    55756: contig of 2285 bp in length
          *    55757    55856: gap of unknown length
          *    55857    58019: contig of 2163 bp in length
          *    58020    58119: gap of unknown length
          *    58120    60327: contig of 2208 bp in length
          *    60328    60427: gap of unknown length
          *    60428    64223: contig of 3796 bp in length
          *    64224    64323: gap of unknown length
          *    64324    67218: contig of 2895 bp in length
          *    67219    67318: gap of unknown length
          *    67319    71797: contig of 4479 bp in length
          *    71798    71897: gap of unknown length
          *    71898    76856: contig of 4959 bp in length
          *    76857    76956: gap of unknown length
          *    76957    82060: contig of 5104 bp in length
          *    82061    82160: gap of unknown length
          *    82161    87557: contig of 5397 bp in length
          *    87558    87657: gap of unknown length
          *    87658    94226: contig of 6569 bp in length
          *    94227    94326: gap of unknown length
          *    94327   101878: contig of 7552 bp in length
          *   101879   101978: gap of unknown length
          *   101979   110629: contig of 8651 bp in length
          *   110630   110729: gap of unknown length
          *   110730   119083: contig of 8354 bp in length
          *   119084   119183: gap of unknown length
          *   119184   128919: contig of 9736 bp in length
          *   128920   129019: gap of unknown length
          *   129020   142791: contig of 13772 bp in length
          *   142792   142891: gap of unknown length
          *   142892   156079: contig of 13188 bp in length
          *   156080   156179: gap of unknown length
          *   156180   183092: contig of 26913 bp in length
          *   183093   183192: gap of unknown length
          *   183193   216441: contig of 33249 bp in length.
FEATURES             Location/Qualifiers
     source          1. .216441
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2257C19"
                     /clone_lib="CalTech human BAC library D"
     gap             1042. .1141
                     /estimated_length=unknown
     gap             2371. .2470
                     /estimated_length=unknown
     gap             3553. .3652
                     /estimated_length=unknown
     gap             4684. .4783
                     /estimated_length=unknown
     gap             6179. .6278
                     /estimated_length=unknown
     gap             7443. .7542
                     /estimated_length=unknown
     gap             8662. .8761
                     /estimated_length=unknown
     gap             10031. .10130
                     /estimated_length=unknown
     gap             11492. .11591
                     /estimated_length=unknown
     gap             12968. .13067
                     /estimated_length=unknown
     gap             14174. .14273
                     /estimated_length=unknown
     gap             15841. .15940
                     /estimated_length=unknown
     gap             17243. .17342
                     /estimated_length=unknown
     gap             18450. .18549
                     /estimated_length=unknown
     gap             19684. .19783
                     /estimated_length=unknown
     gap             21248. .21347
                     /estimated_length=unknown
     gap             22953. .23052
                     /estimated_length=unknown
     gap             24453. .24552
                     /estimated_length=unknown
     gap             25826. .25925
                     /estimated_length=unknown
     gap             27110. .27209
                     /estimated_length=unknown
     gap             28293. .28392
                     /estimated_length=unknown
     gap             29745. .29844

  Query Match            32.8%;  Score 1024.4;  DB 14;  Length 216441;
  Best Local Similarity  98.3%;  Pred. No. 4.6e-124;
  Matches 1119;  Conservative    0;  Mismatches   11;  Indels    8;  Gaps    8;

Qy       1775 CATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGTCACAGTGAC 1834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42031 CATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGTCACAGTGAC 42090

Qy       1835 CCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTTTACAGGGAG 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42091 CCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTTTACAGGGAG 42150

Qy       1895 TCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAA 1954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42151 TCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAA 42210

Qy       1955 AGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTC 2014
             |||||||||||||||||||| |||||||||||||||| |||||||||||||| |||||
Db      42211 AGGTGGGTGGGGTGGAGTGAAG-TTGGAGAGGGACGCTG-TTGGTTCTATGTGG-TGGTC 42267

Qy       2015 TGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTT 2074
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      42268 TGTTTCCCGGACAAGAAAAA-TGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTT 42326

Qy       2075 CAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGA 2134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42327 CAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGA 42386

Qy       2135 CGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGA 2194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42387 CGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGA 42446

Qy       2195 TTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAG 2254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42447 -TTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAG 42505

Qy       2255 GGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAG 2314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42506 GGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAG 42565

Qy       2315 GGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGC 2374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42566 GGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGC 42625

Qy       2375 CACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAG 2434
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      42626 CACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTA-TTTTGTTCAGAACTAG 42684

Qy       2435 ACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATC 2494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42685 ACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATC 42744

Qy       2495 CTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCT 2554
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      42745 CTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGC-TTGGCAGCTCCCT 42803

Qy       2555 GACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTG 2614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      42804 GACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTG 42863

Qy       2615 AGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGC 2674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42864 AGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGC 42923

Qy       2675 TTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAA 2734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42924 TTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAA 42983

Qy       2735 GCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATACCTGCCAATGAGTCAAACAT 2794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      42984 GCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATACCTGCCAATGAGTC-AACAT 43042

Qy       2795 AGTATCTTTATGTAGATACTTAGATTACTAAATATATATTTCATCTACTTCTGAAGTTGA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43043 AGTATCTTTATGTAGATACTTAGATTACTAAATATATATTTCATCTACTTCTGAAGTTGA 43102

Qy       2855 TAGTCTTCCCCCCCCCCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCC 2912
              |||||||||||||||||||||||||||||||||||| ||||      |   |||   ||
Db      43103 TAGTCTTCCCCCCCCCCCCACTTTTTTCTTTTTTGAGACAGGGTCTCCCTCTGTCACC 43160
```

```
RESULT 9
AC151531
LOCUS       AC151531              211799 bp    DNA     linear   ROD 21-JUN-2005
DEFINITION  Mus musculus BAC clone RP24-387D22 from chromosome 7, complete
            sequence.
ACCESSION   AC151531
VERSION     AC151531.2  GI:63004089
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 211799)
  AUTHORS   Zheng,X. and Cotton,M.
  TITLE     The sequence of Mus musculus BAC clone RP24-387D22
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 211799)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 211799)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 211799)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 211799)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On May 3, 2005 this sequence version replaced gi:51972398.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BB0387D22


            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e. phred quality
            >=30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone, fosmid clone or direct clone walk sequence.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to obtain the consensus sequence. The
            assembly was confirmed by restriction digest.
            This finishing standard has slightly changed from the previous
            Human standard. Specifically, standards for regions of low sequence
            complexity (such as dinucleotide repeats and small unit tandem
            repeats) have been relaxed. These regions are very prevalent in the
            mouse genome, and the return on extended finishing efforts is
            minimal.
            If a sequence meets the criteria of the above statement, it needs
            no comments or tags. If the criteria are not met, such as ambiguous
            bases, then the region is duly annotated.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The  BAC Library has been constructed by Pieter de Jong and
            coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
            and/or brain genomic DNA. The clone and detailed information can be
            obtained from Pieter de Jong and coworkers at http://www.chori.org

            This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .211799
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="7"
                     /clone="RP24-387D22"
                     /clone_lib="RPCI-24"
     unsure          11573. .11848
                     /note="Unresolved tandem repeat."
     unsure          15154. .15175
                     /note="Sequence derived from one plasmid subclone."
     unsure          57981. .57992
                     /note="Sequence derived from one plasmid subclone."
     unsure          57987
                     /note="Unresolved bases"
     unsure          100216. .100326
                     /note="Sequence derived from one plasmid subclone."
     unsure          100748. .100942
                     /note="Sequence derived from one plasmid subclone."
     unsure          118186. .118246
                     /note="Sequence derived from one plasmid subclone."
     unsure          203111. .203365
                     /note="Unresolved simple sequence repeat."
ORIGIN

  Query Match             24.1%;  Score 753.8;  DB 9;  Length 211799;
  Best Local Similarity   64.5%;  Pred. No. 5.7e-89;
  Matches 1597;  Conservative    0;  Mismatches  692;  Indels  188;  Gaps   24;

Qy        252 GTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTC 311
              || | |  |   | |||| | ||    ||| || |   | |||||||  |||||||  |
Db      59785 GTCTGCAAAGTAGGAAAAAAAAAAAACCTTGAATTTGTAGGGCGCTAAGGTTCCAGAATTA 59844

Qy        312 GGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCC 371
                ||         || ||||||    |||  | ||||||| ||| || || | |  | |||
Db      59845 TCAT---------GGTAGGGAGAAGAGGGAAAGCCACACTTGATCTACCTGGGCTGGGCC 59895

Qy        372 CCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTG-------- 423
              ||||||||| |||  || ||||  || ||      |  |||     |  ||  |
Db      59896 CCGCCTACCCCTGGCCCGGGGGCGGGGAGGGGGCGGGGGGGGCGGGCGGAGGAAGCCCCG 59955
```





```
              ||    ||| ||  |   ||| || |   |    || || |  |          |   |
Db      61983 TAGAGAGCTGCCCAACTGTCTGACCCCATGTTACAGGGAAGGGCGATGCTGGTATCACTA 62042

Qy       2606 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 2665
              |        | |         ||   ||    | | | |||||||||||||||| |||| ||
Db      62043 GTAACCAGGGGTTGTGAGCCTGGCAGCTGAGTTCAAAATGTAACAGCTTTAATTTGGAAT 62102

Qy       2666 TTAAGAAGCTTTTAAAA 2682
              ||||  |||||| || |
Db      62103 TTAAAGAGCTTTAAACA 62119


RESULT 10
AC166160
LOCUS       AC166160              222531 bp    DNA     linear   HTG 26-JUL-2005
DEFINITION  Mus musculus chromosome 7 clone RP24-86O19, *** SEQUENCING IN
            PROGRESS ***, 14 unordered pieces.
ACCESSION   AC166160
VERSION     AC166160.1  GI:71143373
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 222531)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 222531)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT
            ----------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BB0086O19
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 216535 bases at least Q40
            Consensus quality: 217881 bases at least Q30
            Consensus quality: 218799 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     6578: contig of 6578 bp in length
            *     6579     6678: gap of unknown length
            *     6679    13288: contig of 6610 bp in length
            *    13289    13388: gap of unknown length
            *    13389    20217: contig of 6829 bp in length
            *    20218    20317: gap of unknown length
            *    20318    28619: contig of 8302 bp in length
            *    28620    28719: gap of unknown length
            *    28720    42603: contig of 13884 bp in length
            *    42604    42703: gap of unknown length
            *    42704    59669: contig of 16966 bp in length
            *    59670    59769: gap of unknown length
            *    59770    71649: contig of 11880 bp in length
            *    71650    71749: gap of unknown length
            *    71750    86131: contig of 14382 bp in length
            *    86132    86231: gap of unknown length
            *    86232   107186: contig of 20955 bp in length
            *   107187   107286: gap of unknown length
            *   107287   134561: contig of 27275 bp in length
            *   134562   134661: gap of unknown length
            *   134662   215890: contig of 81229 bp in length
            *   215891   215990: gap of unknown length
            *   215991   217539: contig of 1549 bp in length
            *   217540   217639: gap of unknown length
            *   217640   219158: contig of 1519 bp in length
            *   219159   219258: gap of unknown length
            *   219259   222531: contig of 3273 bp in length.
FEATURES             Location/Qualifiers
     source          1. .222531
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="7"
                     /clone="RP24-86O19"
     misc_feature    1. .6578
                     /note="assembly_name:Contig11"
     gap             6579. .6678
                     /estimated_length=unknown
     misc_feature    6679. .13288
                     /note="assembly_name:Contig12"
     gap             13289. .13388
                     /estimated_length=unknown
     misc_feature    13389. .20217
                     /note="assembly_name:Contig13
                     clone_end:T7
                     vector_side:right"
     gap             20218. .20317
                     /estimated_length=unknown
     misc_feature    20318. .28619
                     /note="assembly_name:Contig14"
     gap             28620. .28719
                     /estimated_length=unknown
     misc_feature    28720. .42603
                     /note="assembly_name:Contig15"
     gap             42604. .42703
                     /estimated_length=unknown
     misc_feature    42704. .59669
                     /note="assembly_name:Contig16"
     gap             59670. .59769
                     /estimated_length=unknown
     misc_feature    59770. .71649
                     /note="assembly_name:Contig17"
     gap             71650. .71749
                     /estimated_length=unknown
     misc_feature    71750. .86131
                     /note="assembly_name:Contig18"
     gap             86132. .86231
                     /estimated_length=unknown
     misc_feature    86232. .107186
                     /note="assembly_name:Contig19"
     gap             107187. .107286
                     /estimated_length=unknown
     misc_feature    107287. .134561
                     /note="assembly_name:Contig20
                     clone_end:SP6
                     vector_side:right"
     gap             134562. .134661
                     /estimated_length=unknown
     misc_feature    134662. .215890
                     /note="assembly_name:Contig21"
     gap             215891. .215990
                     /estimated_length=unknown
     misc_feature    215991. .217539
                     /note="assembly_name:Contig5"
     gap             217540. .217639
                     /estimated_length=unknown
```

```
     misc_feature    217640. .219158
                     /note="assembly_name:Contig8"
     gap             219159. .219258
                     /estimated_length=unknown
     misc_feature    219259. .222531
                     /note="assembly_name:Contig9"
ORIGIN

  Query Match          24.1%;  Score 753.8;  DB 14;  Length 222531;
  Best Local Similarity   64.5%;  Pred. No. 5.6e-89;
  Matches 1597;  Conservative    0;  Mismatches  692;  Indels  188;  Gaps   24;

Qy        252 GTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTC 311
              || | |   |    | ||||  | ||     ||| || |    | |||||||  |||||||| |
Db     179225 GTCTGCAAAGTAGGAAAAAAAAAAACCTTGAATTTGTAGGGCGCTAAGGTTCCAGAATTA 179284

Qy        312 GGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCC 371
               ||          || ||||||    |||  | ||||||  ||| || || | |  | |||
Db     179285 TCAT---------GGTAGGGAGAAGAGGGAAAGCCACACTTGATCTACCTGGGCTGGGCC 179335

Qy        372 CCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTG-------- 423
              |||||||||  |||  || ||||  || ||     |  |||    |  ||  |
Db     179336 CCGCCTACCCCTGGCCCGGGGGCGGGGAGGGGGCGGGGGGGGCGGGCGGAGGAAGCCCCG 179395

Qy        424 --------------------GGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGC 463
                                   |||| ||||||||||||    | || || |||||||| |
Db     179396 GGCCTGGACTGGTGGCTGGAGGCGTGTCCTAGGAAACTGGATCCTGCTGCCCTTGGCCGG 179455

Qy        464 GCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGG 523
              | |||||||    || || ||  |                || ||||  | |  |||
Db     179456 GGCTAAGGCGCGCGGCTCGTCTTGGA--------------GACTTGGGGTCGATCAAGG 179500

Qy        524 ATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGA 583
              ||||||  |  ||| ||  ||||| ||||||||| ||||| |     |  | |||||| |||
Db     179501 ATGGGGTCTTCACTCTGGGCTTCTCAGGGAGCCGCCGTGAAAAAGAAAGTGGTGCAAAGA 179560

Qy        584 CACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAA 643
              ||| ||| ||||| ||||| ||| ||||||   ||||| ||| |   ||||    |||||
Db     179561 CACTACGCGTGACCCGGAGCGCGTCTGGGGGAAATGGAGGAGCGGATGGGGAGAAGCTAA 179620

Qy        644 CGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTC 703
              |  | || || | || ||    | |  | ||||||||| || |||||||| || | ||||
Db     179621 CCAGACTACCCCAGC-CGTGGGGACGCCGGAGGCGCACCTCAAGGGTCCCTGGAGAGCTC 179679

Qy        704 CGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGC 763
               || ||     ||||                 |  ||||| |||| |||   |||| |
Db     179680 TGTTGAGAACAGCGG----------------CGAGGACATGAGGAGCTTTCAGTCCGCCC 179723

Qy        764 GGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGG 823
              || |||||| || ||  ||| | |    |  ||| |  || ||| |
Db     179724 TGCTCAAGGCTGGCATCAGGGCTGGAGGCGACCTACGACGTTGAAG-------------- 179769

Qy        824 CGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCT 883
              |  |   |  |||  | |||||||  || ||| ||||| || ||||||||||||||||
Db     179770 -GTTGCAAGTAGGGTACCGCCGGAGGGCAGCGCCATGGCCAGAGAGGAGTGCAAGGCGCT 179828

Qy        884 GCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGG 943
              |||||||| |||||||||||| || |||||||||| ||| | ||||| || || || ||
Db     179829 GCTGGACGCGCTCAACAAGACCACAGCGTGCTACCATCACTTAGTGCTTACTGTGGGCGG 179888

Qy        944 CTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGA 1003
              ||| |||||| | ||| |||| ||  | |||||||| |||| ||||||||||||| |||
Db     179889 CTCCGCGGACACTCAGGACCTACGCGAAGAGCTGCAGAAGACCCGCCAGAAGGCGCGGGA 179948

Qy       1004 GCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGC 1063
              |||||| ||| ||||| |||||||| |||| |  | ||| || |||||| || | ||  |
Db     179949 GCTGGCCGTGGCCACCGGCGCCCGGTTGACCGTCGCGCTCCGGGACCGGAGCTTAGCTAC 180008

Qy       1064 CGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCT 1123
              ||| |||||||| ||||| ||||||||||||||||||| || ||||||||||||||||||
Db     180009 CGAGGAGCGCGCGGAGTTTGAGCGGCTCTGGGTGGCCTTTTCAGGCTGCCTGGACCTGCT 180068

Qy       1124 GGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCG 1183
```

```
                ||||| |||||||  ||| | || | |||||| ||| | ||||||||||||||||||||
Db     180069 CGAAGCCGACATGCAGCGCGCACTAGCGCTGGGAGCCACCTTCCCGCTGCACGCGCCGCG 180128

Qy       1184 GCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCT 1243
              || || || ||||| || ||| | |||| |||||| |  || || |||||||| ||
Db     180129 CCGGCCACTTGTTCGCACTGGGGTGACGGGCGGCTCCTCTGCTGTAGCTGCGCGCGCCCT 180188

Qy       1244 GAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGA 1303
              |||| |||||||||||||| ||| |||||| |||||||||| ||||| || ||||    |
Db     180189 GAGCGCCCGCAGCCTGCGGCACGAAGCGGAGAGCGACTTCGATGTCGCCGATCTGCCCCA 180248

Qy       1304 GCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAA 1363
              | |||||||||| ||||| |||||||||||||||||| ||| ||||||| ||||| |||||
Db     180249 GTTGGAGCGCGAAGTCCTCCAGGTGGGCGAGATGATTGACGACATGGAAATGAAAGTCAA 180308

Qy       1364 CGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGT 1423
              |||||||||||||||| || || || || || ||||| ||||||| |||||||||    |
Db     180309 CGTGCCCCGCTGGACTGTACAGGCGCGACAAGCGGCAGGCGCCGAACTCCTGTCCGGTGC 180368

Qy       1424 CAGCGCCGGCCCCTCCTCGGTCG------TGTCCTTGCAGGAGCGCGGGGGGGGTTGCGA 1477
              ||| |||||| | ||||| | ||       | ||  | |||||||| |||    |||||
Db     180369 CAGTGCCGGCGCGTCCTCAGCCGGTGGCATTTCGGTAGAGGAGCGCGCGGGCCCCTGCGA 180428

Qy       1478 CCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCT 1537
              |||||| |||||||||||| |||| | ||||| |||||||||||| ||| ||||||||||||
Db     180429 CCCCAGCAAGGCCCTGGCTGCCACCGTTTTCAGCGCCGTGCTGTTGGTGGCTGTGGCCCT 180488

Qy       1538 AGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCT 1597
              ||| | || |||||| | |||| | || || |||| || | |||||  | |    |
Db     180489 CGCCTTATGTGTGGCAAAGCTGAGCTGACTGATATCCAACGGTCGTCCGCTGCCCCTTTT 180548

Qy       1598 CCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGT 1657
              |  ||          | || ||||  | |||| ||  | ||  |||| ||| |||
Db     180549 CTTTC------------CAGGGCATGGGCTTCGGGTTTGAACCGT-TAAGGAGAGAGGT 180594

Qy       1658 CCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTC-CAGTGCACATCTGCCTGGGCA 1716
              | | |    |||||||||  | ||| |  | ||||  ||  ||| ||||||||||||
Db     180595 AGTGAGCCTGAATGTGTGCATATGAACATGATCCTTTCAAAGCCCACCTCTGCCTGGGCA 180654

Qy       1717 GGAC-----ACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCCGGCCGTCA 1771
              ||          | ||||    | |||||||||  ||| |||||||     |||  | ||
Db     180655 AAACGTGTTTTGTTTTTTAACATGCTGGCCCAAGAGGAGTTAACAGCAC-AGGGCGGCCA 180713

Qy       1772 GGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT-CACAG 1830
              || ||||||| |||| |||||  ||||||||||    |||||||||||||| | || |||||
Db     180714 GGCCATTACCTCTAATGTCTGTAGGAGCTTTACCTCCTATTAATAGAAATCGGTCCACAG 180773

Qy       1831 TGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTTTACAG 1890
              |||||| || |||||| ||||                        |||  | ||    ||
Db     180774 TGACCCCAGGTCCCTCTGAGT----------------------TTGATGGGTTAAGACGT 180811

Qy       1891 GGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTT 1950
                 || |||| |||||||| | |          || || | | |||||   |  |||
Db     180812 ATGTCCCAAGTTTGGTGCCCTCACGCTGGTTGGTAGCACTCATTCTGTTTTAGACGAGAC 180871

Qy       1951 GAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTT 2010
              |  ||    | |  ||  || |   | | ||| |  | ||  |     |  | |
Db     180872 AATTTTGTTAAGGTGTGGGGTTGTATCAAGCATAGTGGTGGTGGGGG-----AGGGAGAC 180926

Qy       2011 GGTCTGTTTCCCGGACAAGAAA--AATTGCAATCAAATGTCAGCAGCTTTTATTACCTTA 2068
              | |||| ||  |||||||||  | ||||| | | |||   |  | |||| ||||||||
Db     180927 ATTTTGTTCCCTAGACAAGAAACCAGTTGCAGTAAGATGCGCGTGGGTTTTGTTACCTTA 180986

Qy       2069 ATCTTTCAGGGCCTA-AATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTC 2127
              ||| ||  |||||||    ||| |||| |||| | || ||  | || | ||||| ||||||
Db     180987 ATCCTTTGGGGCCTAGGCTTTGGGAGTGTGTACGGATAGAGGCTCGTGCAAAGTGCTTTC 181046

Qy       2128 TCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTA 2187
              |||  ||  | |||| || |||    ||||||  |  | ||| ||| ||| |||| ||
Db     181047 TCTGTGATACACTACTGCACTTG---GCAGAGCCTGCCTATTAGTCACCAGGAGCTGC-- 181101

Qy       2188 GAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAAC 2247
                                 | ||   |  |    |   ||||| |||||||
```

```
Db     181102 -------------------AGCTATGAGAGGTTTTTGTGTCGTGACCCCATTTAGGGTCA 181142

Qy       2248 CAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGC 2307

Db     181143 CAGCCAAGAAGGTTGTGGTGTATCCTC------------------TGCACACTTGCTTCA 181184

Qy       2308 GGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGG 2367

Db     181185 AGAGAAGACCGCTCCATACTCATCCCCGGATGGGGGAGGGACAGGATGGTGCTGCTTTGG 181244

Qy       2368 GTAGGGCCACCTGG-TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTC 2426

Db     181245 GAAGGGCCACACGGCTATGTAAACAGGCACGTTCTCCTTCTCACCGGCTTACTTTATCCC 181304

Qy       2427 AGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCT-GGGAATCCTCTTTAGAG 2485

Db     181305 AGAATTAGACCACAGTGTTTGACCCTCCTTTGAGGGAGGGCTGGGGAATCCTCTTTGAAG 181364

Qy       2486 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 2545

Db     181365 CTCTTAATCCTATCTAT-CCCTGGAATTTACTCGATAGCCAATAGGAACGCTAG-TTTTC 181422

Qy       2546 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 2605

Db     181423 TAGAGAGCTGCCCAACTGTCTGACCCCATGTTACAGGGAAGGGCGATGCTGGTATCACTA 181482

Qy       2606 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 2665

Db     181483 GTAACCAGGGGTTGTGAGCCTGGCAGCTGAGTTCAAAATGTAACAGCTTTAATTTGGAAT 181542

Qy       2666 TTAAGAAGCTTTTAAAA 2682

Db     181543 TTAAAGAGCTTTAAACA 181559
```

```
RESULT 11
AY229888
LOCUS       AY229888               6560 bp    mRNA    linear   ROD 01-JAN-2004
DEFINITION  Mus musculus RGS9-1 anchoring protein R9AP mRNA, complete cds.
ACCESSION   AY229888
VERSION     AY229888.1  GI:37910119
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6560)
  AUTHORS   Keresztes,G., Mutai,H., Hibino,H., Hudspeth,A.J. and Heller,S.
  TITLE     Expression patterns of the RGS9-1 anchoring protein R9AP in the
            chicken and mouse suggest multiple roles in the nervous system
  JOURNAL   Mol. Cell. Neurosci. 24 (3), 687-695 (2003)
   PUBMED   14664818
REFERENCE   2  (bases 1 to 6560)
  AUTHORS   Keresztes,G. and Heller,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2003) Otolaryngology, Harvard Medical School, 243
            Charles Street, Boston, MA 02114, USA
FEATURES             Location/Qualifiers
     source          1. .6560
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /tissue_type="retina"
     CDS             325. .1038
                     /note="membrane-spanning; SNARE-like"
                     /codon_start=1
                     /product="RGS9-1 anchoring protein R9AP"
                     /protein_id="AAP55847.1"
                     /db_xref="GI:37910120"
                     /translation="MAREECKALLDALNKTTACYHHLVLTVGGSADTQDLREELQKTR
                     QKARELAVATGARLTVALRDRSLATEERAEFERLWVAFSGCLDLLEADMQRALALGAT
```

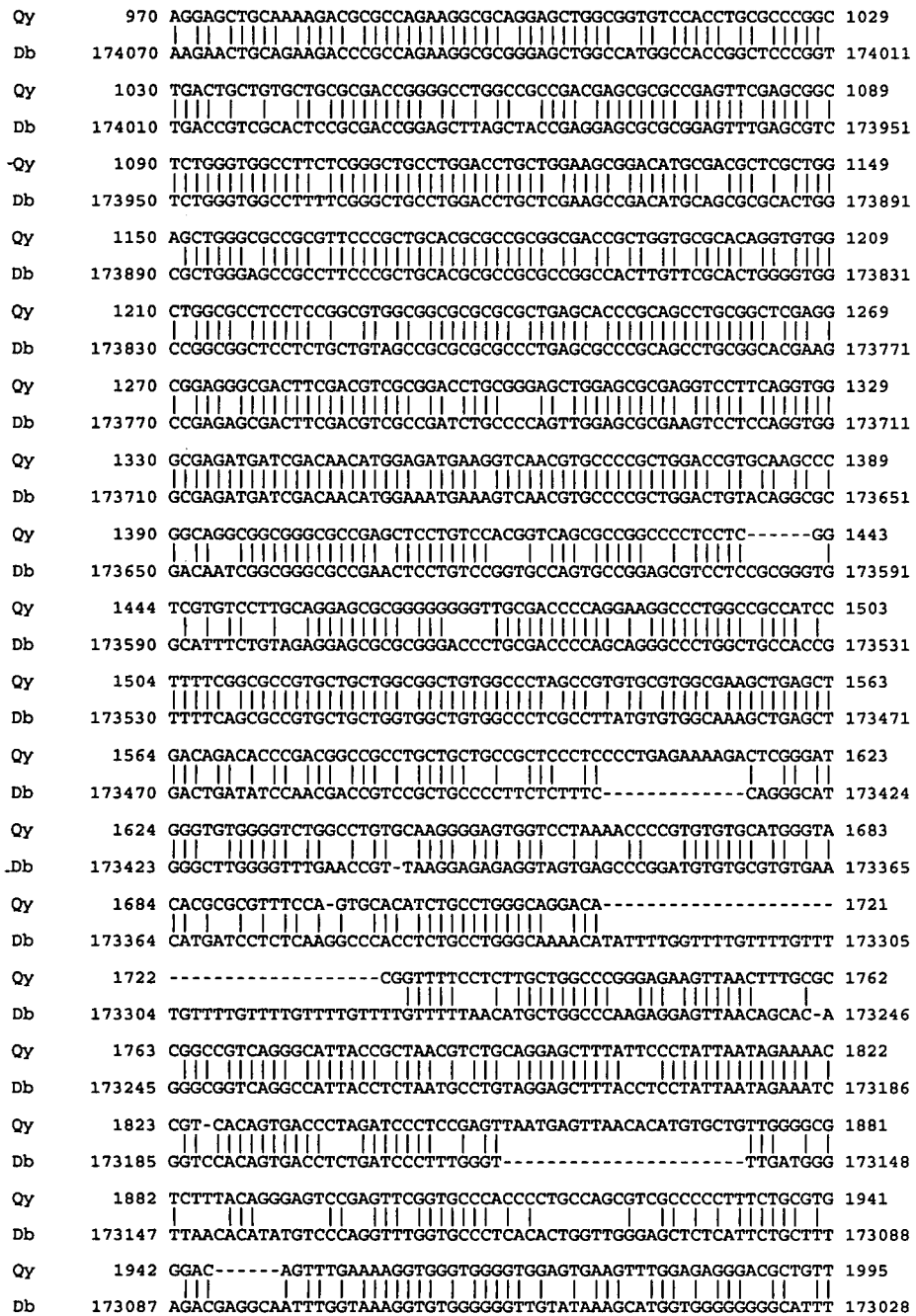
```
                     FPLHAPRRPLVRTGVTGGSSAVAARALSARSLRHEAESDFDVADLPQLEREVLQVGEM
                     IDDMEMKVNVPRWTVQARQAAGAELLSGASAGASSAGGISVEERAGPCDPSKALAATV
                     FSAVLLVAVALALCVAKLS"
ORIGIN

  Query Match             23.6%;  Score 738;  DB 9;  Length 6560;
  Best Local Similarity   66.0%;  Pred. No. 1.3e-86;
  Matches 1450;  Conservative    0;  Mismatches  610;  Indels  136;  Gaps   21;

Qy        505 GAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGA 564

Db          4 GACTTGGGGTCGATCAAGGATGGGGTCTTCACTCTGGGCTTCTCAGGGAGCCGCCGTGAA 63

Qy        565 GGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGA 624

Db         64 AAAGAAAGTGGTGCAAAGACACTACGCGTGACCCGGAGCGCGTCTGGGGGAAATGGAGGA 123

Qy        625 GGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTC 684

Db        124 GCGGATGGGGAGAAGCTAACCAGACTACCCCAGC-CGTGGGGACGCCGGAGGCGCACCTC 182

Qy        685 GAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATG 744

Db        183 AAGGGTCCCTGGAGAGCTCTGTTGAGAACAGCGG----------------CGAGGACATG 226

Qy        745 AAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGC 804

Db        227 AGGAGCTTTCAGTCCGCCCTGCTCAAGGCTGGCATCAGGGCTGGAGGCGACCTACGACGT 286

Qy        805 TGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGA 864

Db        287 TGAAG---------------GTTGCAAGTAGGGTACCGCCGGAGGGCAGCGCCATGGCCA 331

Qy        865 GGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACC 924

Db        332 GAGAGGAGTGCAAGGCGCTGCTGGACGCGCTCAACAAGACCACAGCGTGCTACCATCACT 391

Qy        925 TGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAGA 984

Db        392 TAGTGCTTACTGTGGGCGGCTCCGCGGACACTCAGGACCTACGCGAAGAGCTGCAGAAGA 451

Qy        985 CGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGC 1044

Db        452 CCCGCCAGAAGGCGCGGGAGCTGGCCGTGGCCACCGGCGCCCGGTTGACCGTCGCGCTCC 511

Qy       1045 GCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCT 1104

Db        512 GGGACCGGAGCTTAGCTACCGAGGAGCGCGCGGAGTTTGAGCGGCTCTGGGTGGCCTTTT 571

Qy       1105 CGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGT 1164

Db        572 CAGGCTGCCTGGACCTGCTCGAAGCCGACATGCAGCGCGCACTAGCGCTGGGAGCCACCT 631

Qy       1165 TCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCCG 1224

Db        632 TCCCGCTGCACGCGCCGCGCCGGCCACTTGTTCGCACTGGGGTGACGGGCGGCTCCTCTG 691

Qy       1225 GCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTCG 1284

Db        692 CTGTAGCTGCGCGCGCCCTGAGCGCCCGCAGCCTGCGGCACGAAGCGGAGAGCGACTTCG 751

Qy       1285 ACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGACA 1344

Db        752 ATGTCGCCGATCTGCCCCAGTTGGAGCGCGAAGTCCTCCAGGTGGGCGAGATGATTGACG 811

Qy       1345 ACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGCG 1404

Db        812 ACATGGAAATGAAAGTCAACGTGCCCCGCTGGACTGTACAGGCGCGACAAGCGGCAGGCG 871

Qy       1405 CCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCG------TGTCCTTGCAGG 1458

Db        872 CCGAACTCCTGTCCGGTGCCAGTGCCGGCGCGTCCTCAGCCGGTGGCATTTCGGTAGAGG 931
```

```
Qy       1459 AGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGC 1518
             |||||| |||    |||||||||| |||||||||| |||| | ||||| |||||||||
Db        932 AGCGCGCGGGCCCCTGCGACCCCAGCAAGGCCCTGGCTGCCACCGTTTTCAGCGCCGTGC 991

Qy       1519 TGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACG 1578
             || ||| |||||||||||| ||| | || ||||| |||||||||||| || | || |||
Db        992 TGTTGGTGGCTGTGGCCCTCGCCTTATGTGTGGCAAAGCTGAGCTGACTGATATCCAACG 1051

Qy       1579 GCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGG 1638
             | || | ||||  |  ||  ||          |  || |||||  | |||| ||
Db       1052 GTCGTCCGCTGCCCCTTTTCTTTC-------------CAGGGCATGGGCTTCGGGTTTGA 1098

Qy       1639 CCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTC-CA 1697
              | ||  |||| ||| |||  | |   |    |||||||||  | ||| |  | ||||  |
Db       1099 ACCGTT-AAGGAGAGAGGTAGTGAGCCTGAATGTGTGCATATGAACATGATCCTTTCAAA 1157

Qy       1698 GTGCACATCTGCCTGGGCAGGAC-----ACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTT 1752
              |  ||| |||||||||||||  ||        | ||||    | |||||||||  ||| ||||
Db       1158 GCCCACCTCTGCCTGGGCAAAACGTGTTTTGTTTTTTAACATGCTGGCCCAAGAGGAGTT 1217

Qy       1753 AACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATT 1812
              |||    |  ||| | |||| ||||||| |||| ||||| ||||||||||   ||||||
Db       1218 AACAGCAC-AGGGCGGCCAGGCCATTACCTCTAATGTCTGTAGGAGCTTTACCTCCTATT 1276

Qy       1813 AATAGAAAACCGT-CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGC 1871
              |||||||| | || |||||||||| || |||||| ||||
Db       1277 AATAGAAATCGGTCCACAGTGACCCCAGGTCCCTCTGAGT-------------------- 1316

Qy       1872 TGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCC 1931
               |||  | ||    ||        || |||| ||||||| | |          || || | |
Db       1317 --TTGATGGGTTAAGACGTATGTCCCAAGTTTGGTGCCCTCACGCTGGTTGGTAGCACTC 1374

Qy       1932 TTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGC 1991
              |||||  |  |||    |   ||   | |  ||  || |   | | || |  |
Db       1375 ATTCTGTTTTAGACGAGACAATTTTGTTAAGGTGTGGGGTTGTATCAAGCATAGTGGTGG 1434

Qy       1992 TGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAA--AATTGCAATCAAATGTC 2049
              ||   |     | |  |    | |||| ||  |||||||||  | ||||| | | |||
Db       1435 TGGGGG-----AGGGAGACATTTTGTTCCCTAGACAAGAAACCAGTTGCAGTAAGATGCG 1489

Qy       2050 AGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA-AATTTAGGAGAGTGTCCTGAGAGCA 2108
              |  | |||| |||||||||| ||  |||| ||   ||| |||| |||| | || ||
Db       1490 CGTGGGTTTTGTTACCTTAATCCTTTGGGGCGTAGGCTTTGGGAGTGTGTACGGATAGAG 1549

Qy       2109 GTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCAT 2168
              | || | ||||| |||||||||  ||  | |||| || ||   | ||||||  |  | ||
Db       1550 GCTCGTGCAAAGTGCTTTCTCTGTGATACACTACTGCACT---TGGCAGAGCCTGCCTAT 1606

Qy       2169 TCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGG 2228
              | ||| ||| |||| ||                        | ||   |  |     |  |
Db       1607 TAGTCACCAGGAGCTGC--------------------AGCTATGAGAGGTTTTTGTGTC 1645

Qy       2229 CTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCC 2288
              ||||||  |||||||    ||   | |  || |       ||  |||
Db       1646 GTGACCCCATTTAGGGTCACAGCCAAGAAGGTTGTGGTGTATCCTCT------------- 1692

Qy       2289 AGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGA 2348
                |||||||||| |   || |||  | ||||  |  || ||   ||||||||   ||
Db       1693 -----GCACACTTGCTTCAAGAGAAGACCGCTCCATACTCATCCCCGGATGGGGGAGGGA 1747

Qy       2349 GGAAGAGCAGAGGCTTGGGGTAGGGCCAC-CTGGTGTTTAAACAGGCACTTTCTCCTTCT 2407
              |  |  |||| ||| |||||||| | | | ||||||||||| ||||||||||
Db       1748 CAGGATGGTGCTGCTTTGGGAAGGGCCACACGGCTATGTAAACAGGCACGTTCTCCTTCT 1807

Qy       2408 CTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGC 2467
              |   |||||| |||     ||||| ||||||| ||||||||| |||||||||   |||||||
Db       1808 CACCGGCTTACTTTATCCCAGAATTAGACCACAGTGTTTGACCCTCCTTTGAGGGAGGGC 1867

Qy       2468 T-GGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCA 2526
              | |||||||||||||  ||| |||||||||| ||| |||||||| | |  | | ||||
Db       1868 TGGGGAATCCTCTTTGAAGCTCTTAATCCTATCTAT-CCCTGGAATTTACTCGATAGCCA 1926

Qy       2527 GTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAAT 2586
```

```
                  |||||  ||| | |||   ||    ||| ||  |   ||| || |   |    || ||
Db       1927 ATAGGAACGCTAG-TTTTCTAGAGAGCTGCCCAACTGTCTGACCCCATGTTACAGGGAAG 1985

Qy       2587 GTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGT 2646
              | |          | ||        | ||          ||   ||     | | |||||
Db       1986 GGCGATGCTGGTATCACTAGTAACCAGGGGTTGTGAGCCTGGCAGCTGAGTTCAAAATGT 2045

Qy       2647 AACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAA 2682
              |||||||||| |||| ||||||  |||||| || |
Db       2046 AACAGCTTTAATTTGGAATTTAAAGAGCTTTAAACA 2081
```

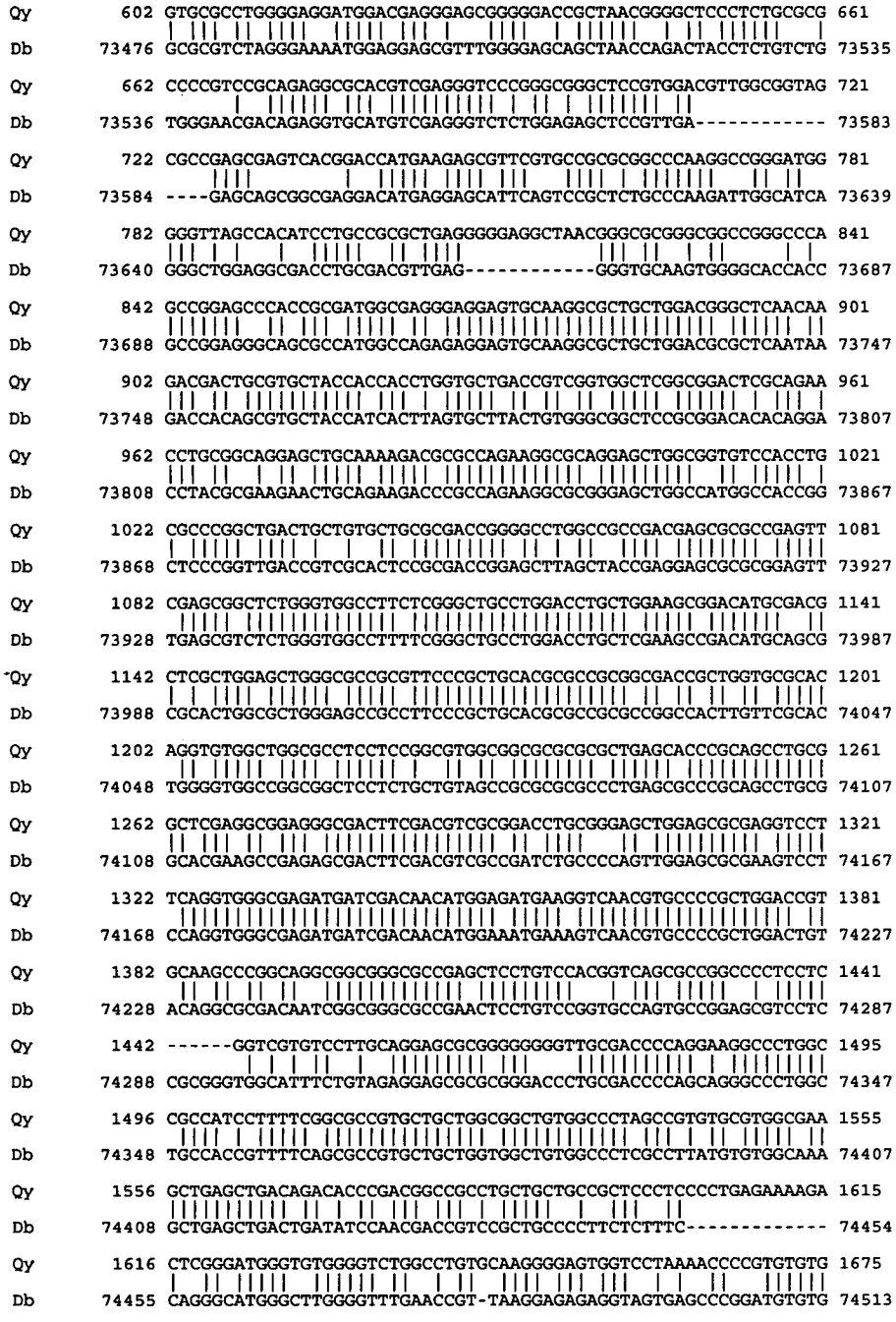
RESULT 12
AC128498/c
LOCUS AC128498 255844 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-328O14, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC128498
VERSION AC128498.3 GI:25138075
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 255844)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von

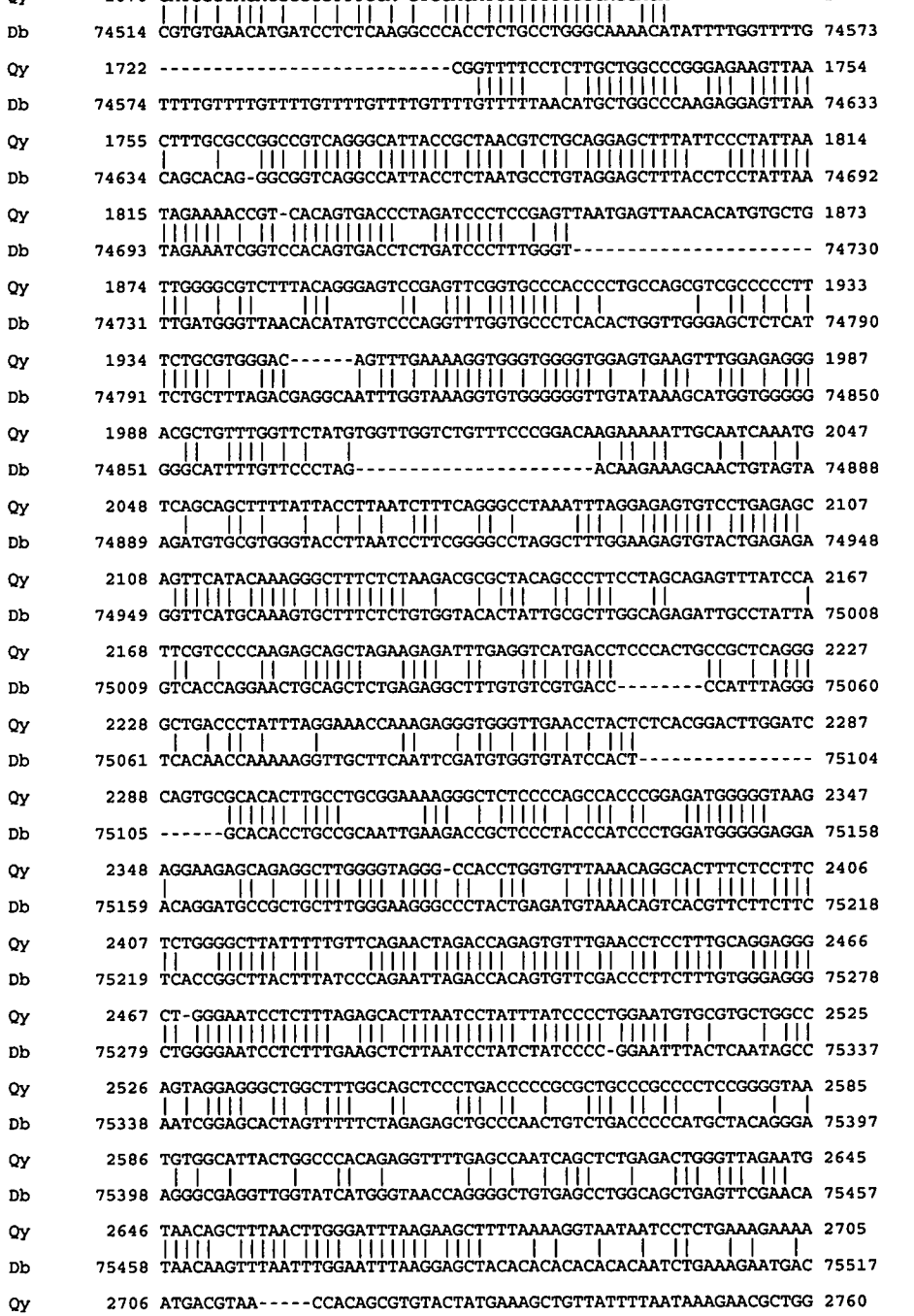
```
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 255844)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 255844)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:23195946.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GYST
                Center clone name: CH230-328O14
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 220111 bases at least Q40
                Consensus quality: 224134 bases at least Q30
                Consensus quality: 227087 bases at least Q20
                Estimated insert size: 227400; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    49930: contig of 49930 bp in length
            *    49931    50030: gap of unknown length
            *    50031   251588: contig of 201558 bp in length
            *   251589   251688: gap of unknown length
            *   251689   255844: contig of 4156 bp in length.
FEATURES            Location/Qualifiers
     source         1. .255844
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-328O14"
     misc_feature   600. .1440
                    /note="clone_boundary
                    clone_end:T7
                    site:
                    end_sequence:RXAON91TJ"
     misc_feature   1490. .4017
                    /note="wgs_contig"
     misc_feature   complement(4520. .5390)
                    /note="clone_boundary
                    clone_end:Sp6
                    site:
                    end_sequence:BZ164083"
     gap            49931. .50030
                    /estimated_length=unknown
     misc_feature   50031. .52743
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   52794. .54086
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   102690. .103774
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   125888. .127459
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   163255. .165112
                    /note="wgs_end_extension
                    clone_end:Sp6"
     gap            251589. .251688
                    /estimated_length=unknown
ORIGIN

  Query Match             22.8%;  Score 711.8;  DB 14;  Length 255844;
  Best Local Similarity   60.7%;  Pred. No. 1.5e-83;
  Matches 1575;  Conservative    0;  Mismatches  817;  Indels  203;  Gaps   20;

Qy        310 TCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAG 369
              | ||   ||  |  |||| |
Db     174686 TGGGGGAGGCGGGGGGGGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 174627

Qy        370 CCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGG 429
                  |  |  | ||| || |||||  | |      ||  || |||  | |
Db     174626 NNNNNNNNNGAGAGGGCGGGAGGGAGGAAGCTCTCGGCCTAGACAGGCGGCTGGAGGCG 174567

Qy        430 TCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGC 489
              |      ||  | | || || |||||||||   | ||||||||    || | |||
Db     174566 TGTCCTAGAAAACGATCCTGCTGCCCTTGGCCAGGGCTAAGGCGCGCGGCTCAGCTTGAA 174507

Qy        490 AGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCA 549
              | ||                |    |||||||||| | |||| ||  | ||| ||
Db     174506 GACTTGGG----------------GTCGATCAGGATGGGGTCTGCACTCTGGGTTTCTCA 174463

Qy        550 GGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCT 609
              ||||||| ||||| |    |  |   |||| |||| | ||| ||||||||||| ||| ||
Db     174462 GGGAGCCGCCGTGAAAAAGAAAGTTATGCAAAGACCCTACGCGTGACTCGGAGCGCGTCT 174403

Qy        610 GGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCC 669
              ||||  ||||| ||| |   ||||  | ||||||  | || ||||||   |      |
Db     174402 AGGGAAAATGGAGGAGCGTTTGGGGAGCAGCTAACCAGACTACCTCTGTCTGTGGGAACG 174343

Qy        670 GCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGC 729
              |||||| ||| |||||||||| | || | ||||||| ||     ||||    |
Db     174342 ACAGAGGTGCATGTCGAGGGTCTCTGGAGAGCTCCGTTGAGAGCAGCGGCGAGGA----- 174288

Qy        730 GAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGC 789
                  ||||| |||| |||   |||| | |||||||   || ||  ||| | |
Db     174287 ----------CATGAGGAGCATTCAGTCCGCTCTGCCCAAGATTGGCATCAGGGCTGGA 174239

Qy        790 CACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGC 849
              |  |||||  || |||||||              ||  | ||     | | |||||||
Db     174238 GGCGACCTGCGACGTTGAGGGGT-----------GCAAGTGGGGCACCACCGCCGGAGG 174191

Qy        850 CCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTG 909
              || ||| ||||| || |||||||||||||||||||||||| |||||| ||||| || |
Db     174190 GCAGCGCCATGGCCAGAGAGGAGTGCAAGGCGCTGCTGGACGCGCTCAATAAGACCACAG 174131

Qy        910 CGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGC 969
              |||||||||| ||| | |||| || || || |||| |||||| | ||| |||| ||
Db     174130 CGTGCTACCATCACTTAGTGCTTACTGTGGGCGGCTCCGCGGACACACAGGACCTACGCG 174071
```



RESULT 13
AC136661
LOCUS AC136661 220118 bp DNA linear HTG 23-NOV-2002
DEFINITION Rattus norvegicus clone CH230-486F10, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC136661
VERSION AC136661.2 GI:25188309
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

```
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 220118)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 220118)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 220118)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 23, 2002 this sequence version replaced gi:24756853.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: KBUP
                Center clone name: CH230-486F10
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 189341 bases at least Q40
                Consensus quality: 190385 bases at least Q30
                Consensus quality: 191057 bases at least Q20
                Estimated insert size: 195411; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   220118: contig of 220118 bp in length.
FEATURES            Location/Qualifiers
     source         1. .220118
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-486F10"
     misc_feature   1. .1070
                    /note="wgs_contig"
     misc_feature   211796. .212627
                    /note="clone_boundary
                    clone_end:T7
                    site:MboI
                    end_sequence:BZ203379"
     misc_feature   215456. .216816
                    /note="wgs_end_extension
                    clone_end:T7"
     misc_feature   217092. .220118
                    /note="wgs_end_extension
                    clone_end:T7"
ORIGIN

  Query Match            22.7%;  Score 709.8;  DB 14;  Length 220118;
  Best Local Similarity  61.9%;  Pred. No. 2.8e-83;
  Matches 1573;  Conservative    0;  Mismatches  767;  Indels  203;  Gaps   20;

Qy        362 GCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGG 421
              ||| || | |  |    | |  |  |  | ||| || |||||  | |      ||  ||
Db      73252 GCGGGGGGGCGGGGGGGCGAGAGGGCGGGAGGGAGGAAGCTCTCGGCCTAGACAGGCGGC 73311

Qy        422 TGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGC 481
              |||  | ||       ||  | | || || |||||||||   | ||||||||    || |
Db      73312 TGGAGGCGTGTCCTAGAAAACGATCCTGCTGCCCTTGGCCAGGGCTAAGGCGCGCGGCTC 73371

Qy        482 GGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCA 541
              |||       | |                |    |    |||||||||  | ||||  ||
Db      73372 AGCTTGAAGACTT---------------GGGGTCGATCAGGATGGGGTCTGCACTCTGG 73415

Qy        542 GCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGA 601
              | ||| |||||||| ||||| |    |  |   |||| |||| | ||| ||||||||||
Db      73416 GTTTCTCAGGGAGCCGCCGTGAAAAAGAAAGTTATGCAAAGACCCTACGCGTGACTCGGA 73475
```


Qy 602 GTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCG 661
Db 73476 GCGCGTCTAGGGAAAATGGAGGAGCGTTTGGGGAGCAGCTAACCAGACTACCTCTGTCTG 73535
Qy 662 CCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAG 721
Db 73536 TGGGAACGACAGAGGTGCATGTCGAGGGTCTCTGGAGAGCTCCGTTGA---------- 73583
Qy 722 CGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGG 781
Db 73584 ----GAGCAGCGGCGAGGACATGAGGAGCATTCAGTCCGCTCTGCCCAAGATTGGCATCA 73639
Qy 782 GGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCA 841
Db 73640 GGGCTGGAGGCGACCTGCGACGTTGAG----------GGGTGCAAGTGGGGCACCACC 73687
Qy 842 GCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAA 901
Db 73688 GCCGGAGGGCAGCGCCATGGCCAGAGAGGAGTGCAAGGCGCTGCTGGACGCGCTCAATAA 73747
Qy 902 GACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAA 961
Db 73748 GACCACAGCGTGCTACCATCACTTAGTGCTTACTGTGGGCGGCTCCGCGGACACACAGGA 73807
Qy 962 CCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTG 1021
Db 73808 CCTACGCGAAGAACTGCAGAAGACCCGCCAGAAGGCGCGGGAGCTGGCCATGGCCACCGG 73867
Qy 1022 CGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTT 1081
Db 73868 CTCCCGGTTGACCGTCGCACTCCGCGACCGGAGCTTAGCTACCGAGGAGCGCGCGGAGTT 73927
Qy 1082 CGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACG 1141
Db 73928 TGAGCGTCTCTGGGTGGCCTTTTCGGGCTGCCTGGACCTGCTCGAAGCCGACATGCAGCG 73987
Qy 1142 CTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCAC 1201
Db 73988 CGCACTGGCGCTGGGAGCCGCCTTCCCGCTGCACGCGCCGCGCCGGCCACTTGTTCGCAC 74047
Qy 1202 AGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCG 1261
Db 74048 TGGGGTGGCCGGCGGCTCCTCTGCTGTAGCCGCGCGCGCCCTGAGCGCCCGCAGCCTGCG 74107
Qy 1262 GCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCT 1321
Db 74108 GCACGAAGCCGAGAGCGACTTCGACGTCGCCGATCTGCCCCAGTTGGAGCGCGAAGTCCT 74167
Qy 1322 TCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGT 1381
Db 74168 CCAGGTGGGCGAGATGATCGACAACATGGAAATGAAAGTCAACGTGCCCCGCTGGACTGT 74227
Qy 1382 GCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTC 1441
Db 74228 ACAGGCGCGACAATCGGCGGGCGCCGAACTCCTGTCCGGTGCCAGTGCCGGAGCGTCCTC 74287
Qy 1442 ------GGTCGTGTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 1495
Db 74288 CGCGGGTGGCATTTCTGTAGAGGAGCGCGCGGGACCCTGCGACCCCAGCAGGGCCCTGGC 74347
Qy 1496 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
Db 74348 TGCCACCGTTTTCAGCGCCGTGCTGCTGGTGGCTGTGGCCCTCGCCTTATGTGTGGCAAA 74407
Qy 1556 GCTGAGCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGA 1615
Db 74408 GCTGAGCTGACTGATATCCAACGACCGTCCGCTGCCCCTTCTCTTTC---------- 74454
Qy 1616 CTCGGGATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTG 1675
Db 74455 CAGGGCATGGGCTTGGGGTTTGAACCGT-TAAGGAGAGAGGTAGTGAGCCCGGATGTGTG 74513


Qy 1676 CATGGGTACACGCGCGTTTCCA-GTGCACATCTGCCTGGGCAGGACA---------- 1721
Db 74514 CGTGTGAACATGATCCTCTCAAGGCCCACCTCTGCCTGGGCAAAACATATTTTGGTTTTG 74573
Qy 1722 ----------CGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAA 1754
Db 74574 TTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTTAACATGCTGGCCCAAGAGGAGTTAA 74633
Qy 1755 CTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAA 1814
Db 74634 CAGCACAG-GGCGGTCAGGCCATTACCTCTAATGCCTGTAGGAGCTTTACCTCCTATTAA 74692
Qy 1815 TAGAAAACCGT-CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTG 1873
Db 74693 TAGAAATCGGTCCACAGTGACCTCTGATCCCTTTGGGT---------- 74730
Qy 1874 TTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTT 1933
Db 74731 TTGATGGGTTAACACATATGTCCCAGGTTTGGTGCCCTCACACTGGTTGGGAGCTCTCAT 74790
Qy 1934 TCTGCGTGGGAC------AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGG 1987
Db 74791 TCTGCTTTAGACGAGGCAATTTGGTAAAGGTGTGGGGGGTTGTATAAAGCATGGTGGGGG 74850
Qy 1988 ACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATG 2047
Db 74851 GGGCATTTTGTTCCCTAG----------ACAAGAAAGCAACTGTAGTA 74888
Qy 2048 TCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGC 2107
Db 74889 AGATGTGCGTGGGTACCTTAATCCTTCGGGGCCTAGGCTTTGGAAGAGTGTACTGAGAGA 74948
Qy 2108 AGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCA 2167
Db 74949 GGTTCATGCAAAGTGCTTTCTCTGTGGTACACTATTGCGCTTGGCAGAGATTGCCTATTA 75008
Qy 2168 TTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGG 2227
Db 75009 GTCACCAGGAACTGCAGCTCTGAGAGGCTTTGTGTCGTGACC--------CCATTTAGGG 75060
Qy 2228 GCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATC 2287
Db 75061 TCACAACCAAAAAGGTTGCTTCAATTCGATGTGGTGTATCCACT---------- 75104
Qy 2288 CAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAG 2347
Db 75105 ------GCACACCTGCCGCAATTGAAGACCGCTCCCTACCCATCCCTGGATGGGGGAGGA 75158
Qy 2348 AGGAAGAGCAGAGGCTTGGGGTAGGG-CCACCTGGTGTTTAAACAGGCACTTTCTCCTTC 2406
Db 75159 ACAGGATGCCGCTGCTTTGGGAAGGGCCCTACTGAGATGTAAACAGTCACGTTCTTCTTC 75218
Qy 2407 TCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGG 2466
Db 75219 TCACCGGCTTACTTTATCCCAGAATTAGACCACAGTGTTCGACCCTTCTTTGTGGGAGGG 75278
Qy 2467 CT-GGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCC 2525
Db 75279 CTGGGGAATCCTCTTTGAAGCTCTTAATCCTATCTATCCCC-GGAATTTACTCAATAGCC 75337
Qy 2526 AGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAA 2585
Db 75338 AATCGGAGCACTAGTTTTTCTAGAGAGCTGCCCAACTGTCTGACCCCCATGCTACAGGGA 75397
Qy 2586 TGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATG 2645
Db 75398 AGGGCGAGGTTGGTATCATGGGTAACCAGGGGCTGTGAGCCTGGCAGCTGAGTTCGAACA 75457
Qy 2646 TAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAA 2705
Db 75458 TAACAAGTTTAATTTGGAATTTAAGGAGCTACACACACACACACAATCTGAAAGAATGAC 75517
Qy 2706 ATGACGTAA-----CCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

```
Db      75518 GTTAGTACATTAGTACCAAAGCTGGGTTGTTTTTGTTCGTTTTTGTATAAAGGAAACTGG 75577

Qy       2761 GCCATGAACTCATACCTGCCAATGA--------GTCAAACATAGTATCTTTATGTAGATA 2812

Db      75578 GCCACAAGCTCATATAGGACTTTTACATTGTACGTGTAGTCTACTACATGTATGTGTATG 75637

Qy       2813 CTTAGATTACTAAATATATATTT 2835

Db      75638 TGTAGTCTACTAAATATGAATCT 75660
```

RESULT 14
CQ748891
LOCUS CQ748891 708 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 34825 from Patent WO02068579.
ACCESSION CQ748891
VERSION CQ748891.1 GI:42376058
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 34825 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .708
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 22.6%; Score 706.4; DB 6; Length 708;
Best Local Similarity 99.9%; Pred. No. 2.6e-82;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917

Db          1 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 60

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977

Db         61 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 120

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037

Db        121 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 180

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097

Db        181 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 240

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157

Db        241 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 300

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217

Db        301 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 360

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277

Db        361 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 420

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337

Db        421 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 480

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397

Db         481 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 540

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457

Db         541 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 600

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517

Db         601 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 660

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565

Db         661 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 708
```

RESULT 15
CQ746021
LOCUS CQ746021 708 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 31955 from Patent WO02068579.
ACCESSION CQ746021
VERSION CQ746021.1 GI:42365754
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 31955 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .708
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 22.6%; Score 704.8; DB 6; Length 708;
Best Local Similarity 99.7%; Pred. No. 4.2e-82;
Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917

Db          1 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 60

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977

Db         61 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 120

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037

Db        121 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 180

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097

Db        181 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 240

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157

Db        241 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGC 300

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217

Db        301 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 360

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 420

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 480

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 540

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 600

Qy       1458 GAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 660

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 708
```

Search completed: May 11, 2006, 05:54:09
Job time : 15360 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 01:07:20 ; Search time 1766 Seconds
(without alignments)
11782.090 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		3122	100.0	3122	12	ADQ62941	Adq62941 Novel hum
2		2308	73.9	2327	13	ADR07462	Adr07462 Full leng
3	c	1009	32.3	1009	4	AAK83424	Aak83424 Human imm
4	c	1009	32.3	1009	4	AAK83423	Aak83423 Human imm
5	c	1005.8	32.2	1009	4	AAK83422	Aak83422 Human imm
6	c	883.2	28.3	973	4	AAK62785	Aak62785 Human imm
7		846.8	27.1	850	5	AAS93728	Aas93728 DNA encod
8	c	726.6	23.3	1061	6	ABQ40781	Abq40781 Oligonucl
9		726.6	23.3	1061	6	ABQ40780	Abq40780 Oligonucl
10	c	706.4	22.6	708	12	ACH87504	Ach87504 Human gen
11	c	594.4	19.0	1061	6	ABQ40778	Abq40778 Oligonucl
12		594.4	19.0	1061	6	ABQ40779	Abq40779 Oligonucl
13	c	524	16.8	524	12	ACH73793	Ach73793 Human gen
14	c	474.4	15.2	476	4	AAK83430	Aak83430 Human imm
15	c	474.4	15.2	476	4	AAK83427	Aak83427 Human imm
16	c	472.8	15.1	476	4	AAK83426	Aak83426 Human imm
17	c	318.6	10.2	1349	5	AAS72508	Aas72508 DNA encod
18		256.2	8.2	432	5	AAS93725	Aas93725 DNA encod
19	c	180.2	5.8	185371	6	ABT10718	Abt10718 Human bre
20		177.2	5.7	24295	12	ADP31706	Adp31706 Human oes
21	c	176.4	5.7	158001	12	ADL17884	Adl17884 Human pho
22		176	5.6	19300	12	ADP74371	Adp74371 Human X c
23		175.6	5.6	3815	12	ADQ85711	Adq85711 Human tum
24	c	175	5.6	283	4	AAK86976	Aak86976 Human imm
25	c	174.8	5.6	22021	11	ACN44670	Acn44670 Human gen
26	c	174.8	5.6	73100	14	ADZ42280	Adz42280 Human end
27		174.8	5.6	99014	6	ABN96931	Abn96931 Gene #342
28		174.8	5.6	109906	6	ABK94411	Abk94411 DNA encod
29		174.8	5.6	109906	12	ADL08112	Adl08112 Human gen
30		174.6	5.6	65454	12	ADN01773	Adn01773 Human hun
31	c	174.4	5.6	110000	14	ADZ45062_01	Continuation (2 of
32		174.2	5.6	110000	6	ABX08336_09	Continuation (10 o
33		174.2	5.6	110000	12	ADJ25985_09	Continuation (10 o
34		174.2	5.6	110000	12	ADN97989_09	Continuation (10 o
35		174.2	5.6	110000	12	ADO50281_09	Continuation (10 o
36		174.2	5.6	110000	14	AEB85185_09	Continuation (10 o
37		174	5.6	13695	10	ADC86452	Adc86452 Human GPC
38		174	5.6	20001	14	ADZ59513	Adz59513 Secondary
39		174	5.6	64958	14	ADZ13534	Adz13534 Human can
40		174	5.6	65387	13	ABD33284	Abd33284 Human can
41		174	5.6	65387	13	ADR67040	Adr67040 Human can
42	c	174	5.6	89014	14	AEB77360	Aeb77360 Human TGF
43	c	174	5.6	168821	11	ACN44262	Acn44262 Human gen
44	c	173.2	5.5	104644	6	ABQ99653	Abq99653 Human MS4
45		173.2	5.5	143306	6	ABK49586	Abk49586 Human tra

ALIGNMENTS

RESULT 1
ADQ62941
ID ADQ62941 standard; cDNA; 3122 BP.
XX
AC ADQ62941;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human cDNA sequence #102.
XX
KW ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
PN EP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
DR WPI; 2004-535376/52.
DR P-PSDB; ADQ65129.
XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 102; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to

CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 3122 BP; 601 A; 891 C; 1019 G; 611 T; 0 U; 0 Other;

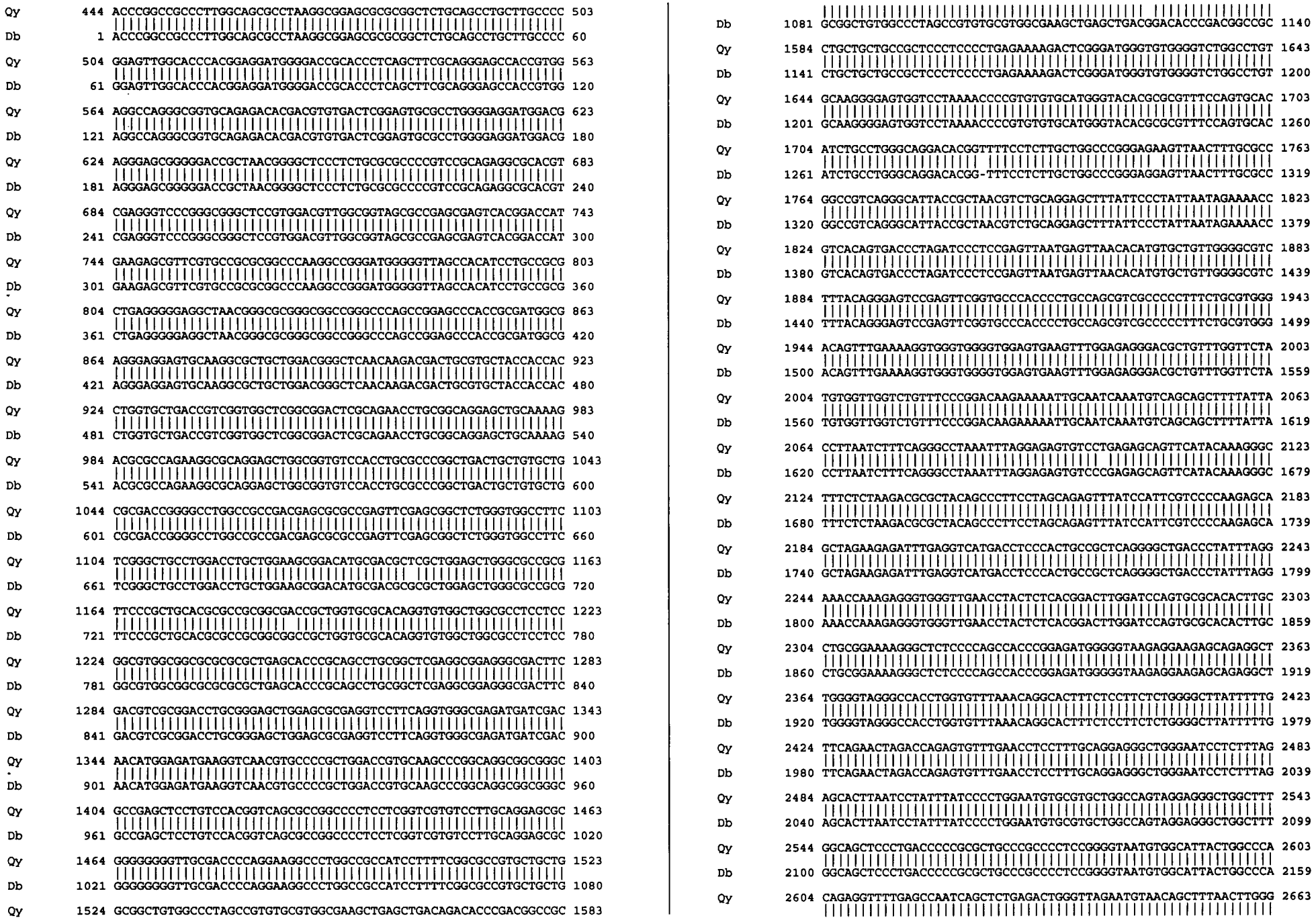
Query Match 100.0%; Score 3122; DB 12; Length 3122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440

Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920

Qy       1921 GCGTCGCCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880

Qy       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940

Qy       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000

Qy       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060

Qy       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120

Qy       3121 TC 3122
                ||
Db       3121 TC 3122
```

RESULT 2
ADR07462
ID ADR07462 standard; cDNA; 2327 BP.
XX
AC ADR07462;
XX
DT 04-NOV-2004 (first entry)
XX
DE Full length human cDNA useful for treating neurological disease Seq 968.
XX
KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-583265/57.
DR P-PSDB; ADR09418.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 968; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polynucleotide is a full
CC length human cDNA sequence of the invention. NOTE: This sequence is not
CC given in the sequence listing of the specification but can be obtained on
CC CD-ROM from the European Patent Office, Vienna Sub-office.
XX
SQ Sequence 2327 BP; 424 A; 667 C; 788 G; 448 T; 0 U; 0 Other;

Query Match 73.9%; Score 2308; DB 13; Length 2327;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;


Qy 444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
Db 1 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 60
Qy 504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
Db 61 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 120
Qy 564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
Db 121 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 180
Qy 624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
Db 181 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 240
Qy 684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
Db 241 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 300
Qy 744 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 803
Db 301 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 360
Qy 804 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 863
Db 361 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 420
Qy 864 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 923
Db 421 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 480
Qy 924 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 983
Db 481 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 540
Qy 984 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 1043
Db 541 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 600
Qy 1044 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
Db 601 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 660
Qy 1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163
Db 661 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGCGCCGCG 720
Qy 1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223
Db 721 TTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 780
Qy 1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283
Db 781 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 840
Qy 1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343
Db 841 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 900
Qy 1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403
Db 901 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 960
Qy 1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1463
Db 961 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1020
Qy 1464 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1523
Db 1021 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1080
Qy 1524 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGC 1583
Db 1081 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACGGACACCCGACGGCCGC 1140
Qy 1584 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1643
Db 1141 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1200
Qy 1644 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1703
Db 1201 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1260
Qy 1704 ATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCC 1763
Db 1261 ATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGCCCGGGAGGAGTTAACTTTGCGCC 1319
Qy 1764 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1823
Db 1320 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1379
Qy 1824 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1883
Db 1380 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1439
Qy 1884 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1943
Db 1440 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1499
Qy 1944 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 2003
Db 1500 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 1559
Qy 2004 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 2063
Db 1560 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 1619
Qy 2064 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGC 2123
Db 1620 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGC 1679
Qy 2124 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 2183
Db 1680 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 1739
Qy 2184 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 2243
Db 1740 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 1799
Qy 2244 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 2303
Db 1800 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 1859
Qy 2304 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 2363
Db 1860 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 1919
Qy 2364 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 2423
Db 1920 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 1979
Qy 2424 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2483
Db 1980 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2039
Qy 2484 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2543
Db 2040 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2099
Qy 2544 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2603
Db 2100 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2159
Qy 2604 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2663

```
Db       2160 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2219

Qy       2664 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2279

Qy       2724 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2771
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2327


RESULT 3
AAK83424/c
ID   AAK83424 standard; DNA; 1009 BP.
XX
AC   AAK83424;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38236.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
```

```
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 38236; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 1009 BP; 260 A; 269 C; 239 G; 241 T; 0 U; 0 Other;

  Query Match           32.3%;  Score 1009;  DB 4;  Length 1009;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-181;
  Matches 1009;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1766 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 1825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 950

Qy       1826 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 1885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 890

Qy       1886 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 830

Qy       1946 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 770

Qy       2006 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 2065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 710

Qy       2066 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTT 2125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTT 650

Qy       2126 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 2185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 590

Qy       2186 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 2245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 530

Qy       2246 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 2305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 470

Qy       2306 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 2365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 410

Qy       2366 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 2425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 350

Qy       2426 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 2485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 290

Qy       2486 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 2545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 230

Qy       2546 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 2605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 170

Qy       2606 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 2665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 110

Qy       2666 TTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 2725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 50

Qy       2726 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 1

RESULT 4
AAK83423/c
ID   AAK83423 standard; DNA; 1009 BP.
XX
AC   AAK83423;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38235.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
```

PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,

```
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 38235; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 1009 BP; 260 A; 269 C; 239 G; 241 T; 0 U; 0 Other;

  Query Match           32.3%;  Score 1009;  DB 4;  Length 1009;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-181;
  Matches 1009;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1766 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 1825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 950

Qy       1826 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 1885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 890

Qy       1886 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 830

Qy       1946 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 770

Qy       2006 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 2065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 710

Qy       2066 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTT 2125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTT 650

Qy       2126 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 2185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 590

Qy       2186 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 2245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 530

Qy       2246 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 2305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 470

Qy       2306 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 2365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 410

Qy       2366 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 2425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 350

Qy       2426 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 2485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 290

Qy       2486 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 2545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 230

Qy       2546 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 2605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 170

Qy       2606 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 2665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 110

Qy       2666 TTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 2725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 50

Qy       2726 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          49 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 1


RESULT 5
AAK83422/c
ID   AAK83422 standard; DNA; 1009 BP.
XX
AC   AAK83422;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38234.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
```

PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 38234; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 1009 BP; 259 A; 269 C; 239 G; 242 T; 0 U; 0 Other;

Query Match 32.2%; Score 1005.8; DB 4; Length 1009;
Best Local Similarity 99.8%; Pred. No. 2.3e-180;
Matches 1007; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1766 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 1825
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db       1009 CCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT 950

Qy       1826 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 1885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTT 890

Qy       1886 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 TACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGAC 830

Qy       1946 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 AGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATG 770

Qy       2006 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 2065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACC 710

Qy       2066 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTT 2125
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        709 TTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGCTT 650

Qy       2126 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 2185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 TCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGC 590

Qy       2186 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 2245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAA 530

Qy       2246 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 2305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCT 470

Qy       2306 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 2365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTG 410

Qy       2366 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 2425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTT 350

Qy       2426 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 2485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 CAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAG 290

Qy       2486 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 2545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGG 230

Qy       2546 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 2605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACA 170

Qy       2606 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 2665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGAT 110

Qy       2666 TTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 2725
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        109 TTAAGAAGCTTTTAAAAGGTAATAATACTCTGAAAGAAAAATGACGTAACCACAGCGTGT 50

Qy       2726 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 1


RESULT 6
AAK62785/c
ID   AAK62785 standard; cDNA; 973 BP.
XX
AC   AAK62785;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:7845.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
```

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR P-PSDB; AAM90004.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Claim 1; SEQ ID NO 7845; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 973 BP; 255 A; 253 C; 226 G; 234 T; 0 U; 5 Other;

Query Match 28.3%; Score 883.2; DB 4; Length 973;
Best Local Similarity 96.6%; Pred. No. 3.2e-157;
Matches 941; Conservative 1; Mismatches 28; Indels 4; Gaps 4;

```
Qy      1803 ATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAA 1862
             || || |||||||||||| | |||||||||| |||| || || |||||||||| | |
Db       973 ATCCCTTATTAATAGAAACCNTTCACAGTGACCTTAGACCCTTCGGAGTTAATGATTAAC 914

Qy      1863 CACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTG-CCAG 1921
                 ||||| | |||||||||||||||||||||||||||||||| || ||||
Db       913 ACATGTGCTGTTTGGGNCNTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCTTGCCCAG 854

Qy      1922 CGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGG 1981
             |||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db       853 CGTCGCCCCCTTTCTGCTTGGGACAGTTTGAAAAGNTGGGTGGGGTGGAGTGAAGTTTGG 794

Qy      1982 AGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAAT 2041
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       793 AGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAAT 734

Qy      2042 CAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCT 2101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       733 CAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCT 674

Qy      2102 GAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTT 2161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       673 GAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTT 614

Qy      2162 TATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGC 2221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       613 TATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGC 554

Qy      2222 TCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACT 2281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       553 TCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACT 494
```

```
Qy       2282 TGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGG 2341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 TGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGG 434

Qy       2342 GGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACA-GGCACTTTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        433 GGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACACGGCACTTTC 374

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 314

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 254

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        253 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCC-G 195

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        194 GGTAATGTGGCATTACTGG-CCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 136

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        135 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCYTTTAAAAGGTAATAATCCTCTGAAA 76

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 16

Qy       2761 GCCATGAACTCATA 2774
             ||||||||||||||
Db          15 GCCATGAACTCATA 2
```

RESULT 7
AAS93728
ID AAS93728 standard; cDNA; 850 BP.
XX
AC AAS93728;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #29532.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG29541.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

. .

XX
PS Claim 1; SEQ ID NO 29532; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 850 BP; 116 A; 278 C; 340 G; 116 T; 0 U; 0 Other;

Query Match 27.1%; Score 846.8; DB 5; Length 850;
Best Local Similarity 99.8%; Pred. No. 2.4e-150;
Matches 848; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 60

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 120

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 180

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 240

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 300

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 360

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 420

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGT 1195
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        421 GCGACGCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGT 480

Qy       1196 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 540

Qy       1256 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 600

Qy       1316 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        601 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 660

Qy       1376 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 1435
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 720

Qy       1436 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 1495
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 780

Qy       1496 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 840

Qy       1556 GCTGAGCTGA 1565
             ||||||||||
Db        841 GCTGAGCTGA 850


RESULT 8
ABQ40781/c
ID   ABQ40781 standard; DNA; 1061 BP.
XX
AC   ABQ40781;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 27372.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 1061 BP; 348 A; 429 C; 137 G; 147 T; 0 U; 0 Other;

  Query Match             23.3%;  Score 726.6;  DB 6;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 1.2e-127;
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554
              | |||   |||||||| |   |||||||||||||| || |   | || |||| ||||||
Db       1061 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 1002

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614
              | ||||||||  |||||||||| ||||| ||||||||||| ||||||||||  ||||||
Db       1001 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 942

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674
              |||||||||||||||||||| || |||||||| |   | ||||||   ||| || |||
Db        941 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 882

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734
              |||| ||||||||||  |||||||| | ||||||||||||||||||| |||||||||
Db        881 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 822

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794
              |||||  |||||||||||||||||| |||||||   |||| ||||||||||||||||  | ||
Db        821 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 762

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854
              || |||| |||||||||| |||||||||||||||| ||||   || |||||   | |
Db        761 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 702

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914
              |||||||||||||||||| |||||| || |||||||| | || ||||||| ||||||
Db        701 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 642

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974
              ||  |  |  ||||| ||| |||||||||| |||||||| ||| ||||  ||||| |||
Db        641 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 582

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034
              || ||||||||||  |||||||| ||||| |||||||||  |  ||||  ||| ||| |
Db        581 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 522

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094
              | |||| ||||||| |||||  ||| || |||||||||| |||||||||||| | |||
Db        521 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 462

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154
              ||||  || ||||| ||  ||||  || |||||||||| |||||||| || ||||| ||
Db        461 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 402

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214
              |||| ||||||  || || ||||| |||||||  || ||||||| | |||||||| ||||
Db        401 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 342

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274
              |  |  | |||||||||||||||||| |||| |  || ||  ||||| ||||||||||
Db        341 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 282

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334
              ||||| |||||||||||||| |||||||| |||||||||||||  || ||||||||||
Db        281 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 222

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394
              |||||||| || |||||||||||||| ||||||   || |||| |||| |||  ||| ||
Db        221 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 162

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454
              |||||||||| |||| |  |||  ||||| |||| |||    |  ||||||||||  |||
```

```
Db        161 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 102

Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514
              ||||||||||||||||||||||||     |||||||    ||| ||  ||  |||||||||

Db        101 TAGGAGCGCGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 42

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||| || |||||| |||||    ||| ||||||||||||||
Db         41 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1
```

```
RESULT 9
ABQ40780
ID   ABQ40780 standard; DNA; 1061 BP.
XX
AC   ABQ40780;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 27371.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 1061 BP; 147 A; 137 C; 429 G; 348 T; 0 U; 0 Other;

  Query Match             23.3%;  Score 726.6;  DB 6;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 1.2e-127;
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554
              | |||   |||||||||  |   |||||||||||||| ||  |   | ||| |||| ||||||
Db          1 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 60

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614
                | ||||||||  |||||||||  |||||  ||||||||||  |||||||||   ||||||
Db         61 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 120

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674
              |||||||||||||||||||||||| || ||||||||  |    | ||||||   ||| || |||
Db        121 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 180

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734
              ||||  ||||||||||   |||||||| | ||||||||||||||||||| |||||||||
Db        181 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 240

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794
              |||||  |||||||||||||||| |||||||   |||| ||||||||||||||  | ||
Db        241 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 300

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854
                || |||| |||||||||||| |||||||||||||||||| ||||    || |||||   | |
Db        301 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 360

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914
              |||||||||||||||||||| |||||| || |||||||| | || |||||| ||||||
Db        361 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 420

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974
              ||  |  |  ||||| ||| |||||||| |||||||| ||| ||||  ||||| |||||
Db        421 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 480

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034
                || ||||||||||  |||||||| ||||| |||||||||  |  ||||  ||| ||| |
Db        481 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 540

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094
              | |||| ||||||| |||||  ||| || ||||||||||| |||||||||||| | |||
Db        541 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 600

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154
              ||||  || ||||| ||  |||| || |||||||||| |||||||||  || ||||| ||
Db        601 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 660

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214
              |||| ||||||  || || ||||| |||||||  || |||||||  | ||||||||  ||||
Db        661 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 720

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274
              |  |  | |||||||||||||||||||| |||| |  || ||  ||||| ||||||||||
Db        721 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 780

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334
              |||||  |||||||||||||||  |||||||| ||||||||||||  || ||||||||||
Db        781 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 840

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394
              |||||||| || |||||||||||||| ||||||   || |||| |||| |||  ||| ||
Db        841 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 900

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454
              |||||||||| |||| |  |||  ||||| |||| |||     |  ||||||||||  |||
Db        901 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 960

Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514
              |||||||||||||||||||||||     |||||||    ||| ||  ||  |||||||||
Db        961 TAGGAGCGCGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 1020

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||| || |||||| |||||   ||| ||||||||||||||
```

```
Db       1021 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1061


RESULT 10
ACH87504/c
ID   ACH87504 standard; DNA; 708 BP.
XX
AC   ACH87504;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #20699.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 20699; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 708 BP; 104 A; 279 C; 231 G; 94 T; 0 U; 0 Other;

  Query Match           22.6%;  Score 706.4;  DB 12;  Length 708;
  Best Local Similarity   99.9%;  Pred. No. 7.4e-124;
  Matches  707;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 649

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 589

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 529

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 469

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 409

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        408 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 349

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 289

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 229

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 169

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 109

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 49

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1


RESULT 11
ABQ40778/c
ID   ABQ40778 standard; DNA; 1061 BP.
XX
AC   ABQ40778;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 27369.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
```

XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1061 BP; 141 A; 137 C; 344 G; 439 T; 0 U; 0 Other;

Query Match 19.0%; Score 594.4; DB 6; Length 1061;
Best Local Similarity 72.5%; Pred. No. 1.1e-102;
Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

RESULT 12
ABQ40779
ID ABQ40779 standard; DNA; 1061 BP.
XX
AC ABQ40779;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 27370.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.

```
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 1061 BP; 439 A; 344 C; 137 G; 141 T; 0 U; 0 Other;

  Query Match           19.0%;  Score 594.4;  DB 6;  Length 1061;
  Best Local Similarity   72.5%;  Pred. No. 1.1e-102;
  Matches  769;  Conservative    0;  Mismatches  291;  Indels    0;  Gaps    0;

Qy        496 CTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGC 555
              ||| ||||| | ||  |||||||| |  ||    |||||||||||| |||||||   | |
Db          2 CTTACCCCGAAATTAACACCCACGAAAAATAAAAACCGCACCCTCAACTTCGCAAAAAAC 61

Qy        556 CACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAG 615
              ||||||  |  |||   || | || | ||||||||| | ||||| | | |||||    |
Db         62 CACCGTAAAAACCAAAACGATACAAAAACACGACGTATAACTCGAAATACGCCTAAAAAA 121

Qy        616 GATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAG 675
              ||  ||||    | ||    ||||||||||   |||||||| ||||||||||||||| |
Db        122 AATAAACGAAAAAACGAAAAACCGCTAACGAAACTCCCTCTACGCGCCCCGTCCGCAAAA 181

Qy        676 GCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCA 735
              |||||||||   |||||  ||  ||||||  |||||  || || |||||| ||| |||
Db        182 ACGCACGTCGAAAATCCCGAACGAACTCCGTAAACGTTAACGATAACGCCGAACGAATCA 241

Qy        736 CGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATC 795
              || ||||| || | ||||||| ||||||| |||||  |||  ||     ||| |||||||
Db        242 CGAACCATAAAAAACGTTCGTACCGCGCGACCCAAAACCGAAATAAAAATTAACCACATC 301

Qy        796 CTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCG 855
              || |||||||  |      |  ||||||  ||||  || |||  |||| ||| | |||||||
Db        302 CTACCGCGCTAAAAAAAAAACTAACGAACGCGAACGACCGAACCCAACCGAAACCCACCG 361

Qy        856 CGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCT 915
              ||||  |||    |  | | |||  |||| ||  |||  |||||||| |||||| ||| ||
Db        362 CGATAACGAAAAAAAAATACAAAACGCTACTAAACGAACTCAACAAAACGACTACGTACT 421

Qy        916 ACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGC 975
              ||||||||||  | || ||||||| |  |||| || ||||||| ||||| || ||  | |
Db        422 ACCACCACCTAATACTAACCGTCGATAACTCGACGAACTCGCAAAACCTACGACAAAAAC 481

Qy        976 TGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTG 1035
              | ||||| |||||||| ||  |||| | ||  || | ||||||| |||||| || |||
Db        482 TACAAAAAACGCGCCAAAAAACGCAAAAACTAACGATATCCACCTACGCCCGACTAACTA 541

Qy       1036 CTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGG 1095
              || | || ||||||||   |||  |||||||||| |||||||| ||||| || ||||
Db        542 CTATACTACGCGACCGAAACCTAACCGCCGACGAACGCGCCGAATTCGAACGACTCTAAA 601

Qy       1096 TGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGG 1155
              |  |||||||  || |||  |||| ||  || || |||| ||||||  ||||  | ||
Db        602 TAACCTTCTCGAACTACCTAAACCTACTAAAAACGAACATACGACGCGCGCTAAAACTAA 661

Qy       1156 GCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCG 1215
              ||||||||||||||| |||||||||| ||||||||  | ||||||  | |  ||  ||
Db        662 ACGCCGCGTTCCCGCTACACGCGCCGCGACGACCGCTAATACGCACAAATATAACTAACG 721

Qy       1216 CCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGG 1275
              ||||||||| |||  || |||||||||| | ||||||| ||| || |||||  || |
Db        722 CCTCCTCCGACGTAACGACGCGCGCGCTAAACACCCGCAACCTACGACTCGAAACGAAAA 781

Qy       1276 GCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGA 1335
              |||||||||||||||| |||| ||  | ||  | |||||  |||||||  |   ||| |
Db        782 ACGACTTCGACGTCGCGAACCTACGAAAACTAAAACGCGAAATCCTTCAAATAAACGAAA 841

Qy       1336 TGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGG 1395
              | ||||||||||  | || ||  |||||||  |||||||  |||||  ||| |||| ||
Db        842 TAATCGACAACATAAAAATAAAAATCAACGTACCCCGCTAAACCGTACAAACCCGACAAA 901

Qy       1396 CGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGC 1455
              || ||  |||||| ||||| |||||| ||| |||||  |||||||||| |||| |||||
Db        902 CGACGAACGCCGAACTCCTATCCACGATCAACGCCGACCCCTCCTCGATCGTATCCTTAC 961

Qy       1456 AGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCG 1515
              |  | ||||        || ||||||||  ||  ||||  |||||||||||||| |||||
Db        962 AAAAACGCGAAAAAAATTACGACCCCAAAAAAACCCTAACCGCCATCCTTTTCGACGCCG 1021

Qy       1516 TGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              | || ||  || || |  ||||| |||| | |||  ||||
Db        1022 TACTACTAACGACTATAACCCTAACCGTATACGTAACGAA 1061


RESULT 13
ACH73793/c
ID   ACH73793 standard; DNA; 524 BP.
XX
AC   ACH73793;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #6988.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
```

```
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 6988; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 524 BP; 61 A; 212 C; 162 G; 89 T; 0 U; 0 Other;

  Query Match             16.8%;  Score 524;  DB 12;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 2e-89;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        484 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 465

Qy        544 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 405

Qy        604 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 345

Qy        664 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 285

Qy        724 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 225

Qy        784 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGC 165

Qy        844 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 105

Qy        904 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 45

Qy        964 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1007
              |||||||||||||||||||||||||||||||||||||||||||
Db         44 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1


RESULT 14
AAK83430/c
ID   AAK83430 standard; DNA; 476 BP.
XX
AC   AAK83430;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38242.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
```

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 38242; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 476 BP; 59 A; 190 C; 157 G; 70 T; 0 U; 0 Other;

Query Match 15.2%; Score 474.4; DB 4; Length 476;
Best Local Similarity 99.8%; Pred. No. 4.5e-80;
Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 417

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 357

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 297
```

```
Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 237

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 177

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 117

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 57

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGC 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db          56 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGC 1
```

RESULT 15
AAK83427/c
ID AAK83427 standard; DNA; 476 BP.
XX
AC AAK83427;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38239.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.

```
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 38239; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 476 BP; 59 A; 190 C; 157 G; 70 T; 0 U; 0 Other;

  Query Match             15.2%;  Score 474.4;  DB 4;  Length 476;
  Best Local Similarity   99.8%;  Pred. No. 4.5e-80;
  Matches  475;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 417

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 357

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 297

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 237

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 177

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 117

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 57

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGC 1191
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         56 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGC 1

Search completed: May 11, 2006, 01:38:08
Job time : 1771 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 01:31:02 ; Search time 11789 Seconds
(without alignments)
12390.302 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	700.6	22.4	3429	4	AK044285	AK044285 Mus muscu
	2	668.4	21.4	677	6	CF145448	CF145448 UI-HF-CB0
	3	662.4	21.2	670	6	CF145408	CF145408 UI-HF-CB0
	4	649.4	20.8	655	3	BQ187216	BQ187216 UI-E-EJ1-
c	5	646.4	20.7	685	3	BM671616	BM671616 UI-E-CQ1-
	6	646	20.7	660	5	BX116028	BX116028 BX116028
	7	577.4	18.5	972	5	BQ682843	BQ682843 AGENCOURT
	8	560	17.9	578	3	BM707056	BM707056 UI-E-CR1-
	9	534.8	17.1	595	8	W81663	W81663 zd85e06.r1
	10	500	16.0	500	3	BM710194	BM710194 UI-E-CQ1-
c	11	472.8	15.1	476	1	AW302149	AW302149 xs01f06.x
c	12	461.8	14.8	480	3	BM673230	BM673230 UI-E-CR1-
	13	445.2	14.3	1376	3	BM562937	BM562937 AGENCOURT
c	14	431.4	13.8	455	1	AI375213	AI375213 tc10f06.x
c	15	357.2	11.4	792	5	BX919953	BX919953 BX919953
c	16	347.6	11.1	677	5	BX665878	BX665878 BX665878
	17	337	10.8	561	5	BX921379	BX921379 BX921379
c	18	335.2	10.7	752	5	BX667653	BX667653 BX667653
c	19	334.8	10.7	605	3	BI458731	BI458731 603199879
	20	327.4	10.5	553	5	BX915965	BX915965 BX915965
c	21	321.6	10.3	443	9	BZ884827	BZ884827 CH240_283
c	22	313.4	10.0	727	5	BX918257	BX918257 BX918257
	23	289.6	9.3	940	5	BU506122	BU506122 AGENCOURT
	24	287.6	9.2	672	2	BB591662	BB591662 BB591662
	25	266.4	8.5	1128	8	DN684685	DN684685 CGX37-D05
	26	266.4	8.5	1313	8	DN690837	DN690837 CGX72-H01
	27	266.2	8.5	1154	8	DN683797	DN683797 CGX32-B03
c	28	263.2	8.4	449	8	W81222	W81222 zd85e06.s1
	29	255.6	8.2	1101	8	DN687616	DN687616 CGX54-A11
	30	255.2	8.2	718	8	DN878715	DN878715 nae22d04.
	31	242	7.8	677	8	DN876083	DN876083 nae03h03.
c	32	242	7.8	716	8	DN876084	DN876084 nae03h03.
	33	230.6	7.4	701	3	BJ735612	BJ735612 BJ735612
	34	211.4	6.8	755	11	CR867967	CR867967 Sus scrof
	35	211.4	6.8	800	8	CX937965	CX937965 JGI_CAAO5
	36	208.8	6.7	728	5	BX645932	BX645932 DKFZp781B
*	37	206.4	6.6	585	10	CE444344	CE444344 tigr-gss-
	38	197.6	6.3	1077	3	BI730314	BI730314 603350295
	39	196	6.3	1307	8	DN705253	DN705253 CLJ58-C12
c	40	179.8	5.8	588	9	AQ636023	AQ636023 RPCI-11-4
c	41	179.2	5.7	556	7	CR548736	CR548736 DKFZp469E
	42	176.2	5.6	754	5	BX457023	BX457023 BX457023
	43	175	5.6	1123	5	BX404721	BX404721 BX404721
	44	174.8	5.6	564	6	CD694516	CD694516 EST11039
	45	174.6	5.6	1035	2	BF672721	BF672721 602152030

ALIGNMENTS

RESULT 1
AK044285
LOCUS AK044285 3429 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930005K04 product:unknown EST, full insert sequence.
ACCESSION AK044285
VERSION AK044285.1 GI:26090272
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

```
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3429)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Retina RNA was provided by Dr. Stefano Gustincich (Department of
            Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
            02115, USA) whose assistance is gratefully acknowledged. Please
            visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .3429
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A930005K04"
                     /db_xref="taxon:10090"
                     /clone="A930005K04"
                     /tissue_type="retina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1. .3429
                     /note="unknown EST (GB|BI730314, evidence: BLASTN, 99%,
                     match=315)"
ORIGIN

  Query Match             22.4%;  Score 700.6;  DB 4;  Length 3429;
  Best Local Similarity   65.5%;  Pred. No. 9e-112;
  Matches 1440;  Conservative    0;  Mismatches  619;  Indels  138;  Gaps   23;

Qy        505 GAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGA 564
              || ||||    | |   ||||||||| |  ||| ||  |||||| ||||||||| |||||
Db         43 GACTTGGGGTCGATCAAGGATGGGGTCTTCACTCTGGGCTTCTCAGGGAGCCGCCGTGAA 102

Qy        565 GGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGA 624
                |  | |||| | |||||| ||| ||||| |||| ||| ||||||    ||||| ||
Db        103 AAAGAAAGTGGTGAAAAGACACTACGCGTGACCCGGAGCGCGTCTGGGGGAAATGGAGGA 162

Qy        625 GGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTC 684
              | |   ||||    ||||||  | || || | || ||    | |  | ||||||||| ||
Db        163 GCGGATGGGGAGAAGCTAACCAGACTACCCCAGC-CGTGGGGACGCCGGAGGCGCACCTC 221

Qy        685 GAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATG 744
              |||||||||  || | |||| || ||      ||||                 |  ||||
Db        222 GAGGGTCCCTGGAGAGCTCTGTTGAGAACAGCGG----------------CGAGGACATG 265

Qy        745 AAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGC 804
              | |||| |||   |||| | || |||||| || ||  ||| | |   |  ||| |  ||

Db        266 AGGAGCTTTCAGTCCGCCCTGCTCAAGGCTGGCATCAGGGCTGGAGGCGACCTACGACGT 325

Qy        805 TGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGA 864
              ||| |                 | |   |  |||  | |||||||  || ||| ||||| |
Db        326 TGAAG--------------GTTGCAAGTGGGGTACCGCCGGAGGGCAGCGCCATGGCCA 370

Qy        865 GGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACC 924
              | |||||||||||||||||||||||| |||||||||||  || ||||||||||  |||
Db        371 GAGAGGAGTGCAAGGCGCTGCTGGACGCGCTCAACAAGACCACAGCGTGCTACCATCACT 430

Qy        925 TGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAGA 984
              | ||||| || || || ||||| |||||| | ||| |||| ||  | ||||||||  ||||
Db        431 TAGTGCTTACTGTGGGCGGCTCCGCGGACACTCAGGACCTACGCGAAGAGCTGCAGAAGA 490

Qy        985 CGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGC 1044
              | |||||||||||| |||||||| ||| ||||| |||||||| |||| |  | ||| |
Db        491 CCCGCCAGAAGGCGCGGGAGCTGGCCGTGGCCACCGGCGCCCGGTTGACCGTCGCGCTCC 550

Qy       1045 GCG-ACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
              | | ||||| || | ||  |||| |||||||| || || |||||||| ||||||||||
Db        551 GGGCACCGGAGCTTAGCTACCGAGGAGCGCGCGGACTTTGAGCGGCTATGGGTGGCCTTT 610

Qy       1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163
              || ||||| |||||||||||| ||| | |||||||  ||| | || | |||||| ||| |
Db        611 TCAGGCTGGCTGGACCTGCTCGAACCCGACATGCAGCGCGCACTAGCGCTGGGAGCCACC 670

Qy       1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223
              | |||||||||||||| | || || || ||||| || ||| | | ||  | ||
Db        671 CTTCCGCTGCACGCGCCGGGCCGGCCACTTGTTCGCACTGGGGTGACGGCCGGAT-CTAT 729

Qy       1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283
              |  || || |||||||| |||||| |||||||||||||| ||| |||||| ||||||||
Db        730 GCTGTAGCTGCGCGCGCCCTGAGCGCCCGCAGCCTGCGGCACGAAGCGGAGAGCGACTTC 789

Qy       1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343
              || ||||| || ||||    || |||||||||| ||||| ||||||||||||||||||||
Db        790 GATGTCGCCGATCTGCCCCAGTTGGAGCGCGAAGTCCTCCAGGTGGGCGAGATGATCGAC 849

Qy       1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403
              |||||||  ||||| |||||||||||||||||||||| || || || || || | | |
Db        850 GACATGGAAATGAAAGTCAACGTGCCCCGCTGGACTGTACAGGCGCGACAAGCGGCAGAC 909

Qy       1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCG------TGTCCTTGCAG 1457
              |||||  |||||||||    | |||  ||||||  | |||||  | ||    | ||  | ||
Db        910 GCCGAACTCCTGTCCGGTGCCAGTGCCGGCGCGTCCTCAGCCGGTGGCATTTCGGTAGAG 969

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
              |||||||  |||     |||||||||||  |||||||||||  |||| | ||||| ||||||||
Db        970 GAGCGCGCGGGCCCCTGCGACCCCAGCAAGGCCCTGGCTGCCACCGTTTTCAGCGCCGTG 1029

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGAC 1577
              ||| ||| |||||||||||| ||| | || ||||| ||||||||||||| || | || ||
Db       1030 CTGTTGGTGGCTGTGGCCCTCGCCTTATGTGTGGCAAAGCTGAGCTGACTGATATCCAAC 1089

Qy       1578 GGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTG 1637
              || || | ||||  |  ||  ||          | || |||||  | |||| ||
Db       1090 GGTCGTCCGCTGCCCCTTTTCTTTC----------CAGGGCATGGGCTTCGGGTTTG 1136

Qy       1638 GCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTC-C 1696
                | ||  |||| ||| |||  | |  |     |||||||||  | |||  |  ||||
Db       1137 AACCGTT-AAGGAGAGAGGTAGTGAGCCTGAATGTGTGCATATGAACATGATCCTTTCAA 1195

Qy       1697 AGTGCACATCTGCCTGGGCAGGAC-----ACGGTTTTCCTCTTGCTGGCCCGGGAGAAGT 1751
              ||  ||| |||||||||||||  ||        |||||  |||||||||  ||| |||
Db       1196 AGCCCACCTCTGCCTGGGCAAAACGTGTTTTGTTTTTTAACATGCTGGCCCAAGAGGAGT 1255

Qy       1752 TAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTAT 1811
              ||||    |  ||| | |||| ||||||| |||| ||||| ||||||||||    |||||
Db       1256 TAACAGCAC-AGGGCGGCCAGGCCATTACCTCTAATGTCTGTAGGAGCTTTACCTCCTAT 1314

Qy       1812 TAATAGAAAACCGT-CACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTG 1870
              ||||||||| | || |||||||||||  || |||||  ||||
Db       1315 TAATAGAAATCGGTCCACAGTGACCCCAGGTCCCTCTGAGT------------------- 1355
```

```
Qy       1871 CTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCC 1930
Db       1356 ---TTGATGGGTTAAGACGTATGTCCCAAGTTTGGTGCCCTCACGCTGGTTGGTAGCACT 1412

Qy       1931 CTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACG 1990
Db       1413 CATTCTGTTTTAGACGAGACAATTTTGTTAAGGTGTGGGGTTGTATCAAGCATAGTGGTG 1472

Qy       1991 CTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAA--AATTGCAATCAAATGT 2048
Db       1473 GTGGGGG-----AGGGAGACATTTTGTTCCCTAGACAAGAAACCAGTTGCAGTAAGATGC 1527

Qy       2049 CAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA-AATTTAGGAGAGTGTCCTGAGAGC 2107
Db       1528 GCGTGGGTTTTGTTACCTTAATCCTTTGGGGCCTAGGCTTTGGGAGTGTGTACGGATAGA 1587

Qy       2108 AGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCA 2167
Db       1588 GGCTCGTGCAAAGTGCTTTCTCTGTGATACACTACTGCACT---TGGCAGAGCCTGCCTA 1644

Qy       2168 TTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGG 2227
Db       1645 TTAGTCACCAGGAGCTGC--------------------AGCTATGAGAGGTTTTTGTGT 1683

Qy       2228 GCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATC 2287
Db       1684 CGTGACCCCATTTAGGGTCACAGCCAAGAAGGTTGTGGTGTATCCTCT------------ 1731

Qy       2288 CAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAG 2347
Db       1732 ------GCACACTTGCTTCAAGAGAAGACCGCTCCATACTCATCCCCGGATGGGGGAGGG 1785

Qy       2348 AGGAAGAGCAGAGGCTTGGGGTAGGGCCAC-CTGGTGTTTAAACAGGCACTTTCTCCTTC 2406
Db       1786 ACAGGATGGTGCTGCTTTGGGAAGGGCCACACGGCTATGTAAACAGGCACGTTCTCCTTC 1845

Qy       2407 TCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGG 2466
Db       1846 TCACCGGCTTACTTTATCCCAGAATTAGACCACAGTGTTTGACCCTCCTTTGAGGGAGGG 1905

Qy       2467 CT-GGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCC 2525
Db       1906 CTGGGGAATCCTCTTTGAAGCTCTTAATCCTATCTAT-CCCTGGAATTTACTCGATAGCC 1964

Qy       2526 AGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAA 2585
Db       1965 AATAGGAACGCTAG-TTTTCTAGAGAGCTGCCCAACTGTCTGACCCCATGTTACAGGGAA 2023

Qy       2586 TGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATG 2645
Db       2024 GGGCGATGCTGGTATCACTAGTAACCAGGGGTTGTGAGCCTGGCAGCTGAGTTCAAAATG 2083

Qy       2646 TAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAA 2682
Db       2084 TAACAGCTTTAATTTGGAATTTAAAGAGCTTTAAACA 2120
```

```
RESULT 2
CF145448
LOCUS       CF145448                677 bp    mRNA    linear   EST 06-AUG-2003
DEFINITION  UI-HF-CB0-asp-a-08-0-UI.r1 NIH_MGC_210 Homo sapiens cDNA clone
            IMAGE:30569191 5', mRNA sequence.
ACCESSION   CF145448
VERSION     CF145448.1  GI:33260892
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Tim Ratlift
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDNA
            sequence: 129-236, >MIR#SINE/MIR
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30569191"
                     /tissue_type="CNCAP(3)T-225 cell line"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_210"
                     /note="Organ: Prostate; Vector: pT7T3 Pac; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pT7T3 Pac vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CCCAC. Tissue was provided by Tim Ratlift."
ORIGIN
```

Query Match 21.4%; Score 668.4; DB 6; Length 677;
Best Local Similarity 99.7%; Pred. No. 4.7e-106;
Matches 669; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         43 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 102
Db          7 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 66

Qy        103 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 162
Db         67 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 126

Qy        163 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 222
Db        127 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 186

Qy        223 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 282
Db        187 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 246

Qy        283 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 342
Db        247 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 306

Qy        343 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 402
Db        307 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 366

Qy        403 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 462
Db        367 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 426
```

```
Qy        463 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCCGGAGTTGGCACCCACGGAG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCCGGAGTTGGCACCCACGGAG 486

Qy        523 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 546

Qy        583 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTA 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        547 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGAACGCTA 606

Qy        643 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCT 702
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        607 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCNCGGGCGGGCT 666

Qy        703 CCGTGGACGTT 713
              |||||||||||
Db        667 CCGTGGACGTT 677
```

RESULT 3
CF145408

```
LOCUS       CF145408                670 bp    mRNA    linear   EST 06-AUG-2003
DEFINITION  UI-HF-CB0-aso-e-10-0-UI.r1 NIH_MGC_210 Homo sapiens cDNA clone
            IMAGE:30569145 5', mRNA sequence.
ACCESSION   CF145408
VERSION     CF145408.1  GI:33260852
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 670)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Tim Ratlift
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDNA
            sequence: 129-236, >MIR#SINE/MIR
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .670
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30569145"
                     /tissue_type="CNCAP(3)T-225 cell line"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_210"
                     /note="Organ: Prostate; Vector: pT7T3 Pac; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pT7T3 Pac vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CCCAC. Tissue was provided by Tim Ratlift."
ORIGIN
```

```
  Query Match             21.2%;  Score 662.4;  DB 6;  Length 670;
  Best Local Similarity   99.8%;  Pred. No. 5.2e-105;
  Matches  663;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         43 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 66

Qy        103 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 126

Qy        163 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 186

Qy        223 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 246

Qy        283 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 306

Qy        343 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 366

Qy        403 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 426

Qy        463 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 486

Qy        523 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 546

Qy        583 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTA 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        547 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGATGGAGCGGGGGACCGCTA 606

Qy        643 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCT 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCT 666

Qy        703 CCGT 706
              ||||
Db        667 CCGT 670
```

RESULT 4
BQ187216

```
LOCUS       BQ187216                655 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-E-EJ1-ajz-e-14-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajz-e-14-0-UI 5', mRNA sequence.
ACCESSION   BQ187216
VERSION     BQ187216.1  GI:20362767
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
```

```
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 296-333, >GC_rich#Low_complexity (matched compliment)
            Seq primer: M13 REVERSE.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajz-e-14-0-UI"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EJ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes,
                     AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
                     AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
                     Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
                     library was created for the program, Gene Discovery in the
                     Visual System, supported by National Eye Institute (NEI)."
ORIGIN

  Query Match             20.8%;  Score 649.4;  DB 3;  Length 655;
  Best Local Similarity   99.4%;  Pred. No. 9.5e-103;
  Matches  650;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        526 GGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACA 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          2 GGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCANAGACA 61

Qy        586 CGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACG 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACG 121

Qy        646 GGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCG 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCG 181

Qy        706 TGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGG 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGG 241

Qy        766 CCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCG 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCG 301

Qy        826 CGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGC 885
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        302 CGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCNAGGGAGGAGTGCAAGGCGCTTC 361

Qy        886 TGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCT 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCT 421

Qy        946 CGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGC 1005
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        422 CGGCGGACTCGCANAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGC 481

Qy       1006 TGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCG 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCG 541

Qy       1066 ACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGG 601

Qy       1126 AAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGC 1179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGC 655


RESULT 5
BM671616/c
LOCUS       BM671616                685 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CQ1-agc-l-04-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
            UI-E-CQ1-agc-l-04-0-UI 3', mRNA sequence.
ACCESSION   BM671616
VERSION     BM671616.1  GI:18981514
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CQ1-agc-l-04-0-UI"
                     /tissue_type="optic nerve"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CQ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
```

modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-CQ1 is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCATTAAGTG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human optic nerve
TAG_LIB=UI-E-CQ1
TAG_SEQ=CCATTAAGTG"

ORIGIN

Query Match 20.7%; Score 646.4; DB 3; Length 685;
Best Local Similarity 99.6%; Pred. No. 3.1e-102;
Matches 669; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy       2103 AGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTT 2162
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        685 AGAGCAGTTCATACAAA-GGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTT 627

Qy       2163 ATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCT 2222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 ATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCT 567

Qy       2223 CAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTT 2282
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        566 CAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGG-TGAACCTACTCTCACGGACTT 508

Qy       2283 GGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGG 2342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGG 448

Qy       2343 GTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTC 2402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTC 388

Qy       2403 CTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGG 2462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGG 328

Qy       2463 AGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTG 2522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 AGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTG 268

Qy       2523 GCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGG 2582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGG 208

Qy       2583 TAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGA 2642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGA 148

Qy       2643 ATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGA 2702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGA 88

Qy       2703 AAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGC 2762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          87 AAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGC 28

Qy       2763 CATGAACTCATA 2774
             |||||||||| |
Db          27 CATGAACTCAAA 16
```

RESULT 6
BX116028
LOCUS BX116028 660 bp mRNA linear EST 07-FEB-2003
DEFINITION BX116028 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGp998C11795 ; IMAGE:347458, mRNA sequence.
ACCESSION BX116028
VERSION BX116028.1 GI:27880354
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998C11795.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD; contact RZPD (clone@rzpd.de) for further information. Seq primer: M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES Location/Qualifiers
source 1. .660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998C11795 ; IMAGE:347458"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_heart_NbHH19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ORIGIN

Query Match 20.7%; Score 646; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-102;
Matches 646; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2129 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 60

Qy       2189 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 2248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 120

Qy       2249 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 180
```

```
Qy       2309 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 2368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 240

Qy       2369 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 2428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 300

Qy       2429 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 2488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 360

Qy       2489 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 2548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 420

Qy       2549 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 2608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 480

Qy       2609 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 2668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 540

Qy       2669 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 2728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 600

Qy       2729 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 646
```

RESULT 7
BQ682843
LOCUS BQ682843 972 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8495014 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6295149 5', mRNA sequence.
ACCESSION BQ682843
VERSION BQ682843.1 GI:21795522
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 972)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2500 row: o column: 22
High quality sequence stop: 463.
FEATURES Location/Qualifiers
source 1. .972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6295149"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 18.5%; Score 577.4; DB 5; Length 972;
Best Local Similarity 96.0%; Pred. No. 3e-90;
Matches 656; Conservative 0; Mismatches 21; Indels 6; Gaps 6;

```
Qy       1899 AGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGT 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGT 60

Qy       1959 GGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTT 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTT 120

Qy       2019 TCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGG 180

Qy       2079 GCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCG 240

Qy       2139 CTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTG 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTG 300

Qy       2199 AGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTG 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTG 360

Qy       2259 GGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCT 2318
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| ||
Db        361 GGTTGAACCTACTCTCACGGACTTGGATCCAGTGCACACACTTGCCTGCGGAAAAGGACT 420

Qy       2319 CTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCA-C 2377
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |
Db        421 CTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCACAGGCTTGGGGTAGGGCCACC 480

Qy       2378 CTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACC 2437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACC 540

Qy       2438 AGAGTGTTTGAA-CCTCCTTTGCAGGAGGGCTGGG-AATCCTCTTTAGAGCA-CTTAATC 2494
             |||||||||||| |||||||||||||||||||||| |||||||||| | || |||||||
Db        541 AGAGTGTTTGAACCCTCCTTTGCAGGAGGGCTGGGAAATCCTCTTTAAAACACCTTAATC 600

Qy       2495 CTATTTAT-CCCCTGGAATGTGCGTGCTGGCCAGTAGG-AGGGCTGGCTTTGGCAGCTCC 2552
             |||||||| |||||||||||||| ||||||||| ||| |||||||| ||||| ||||||
Db        601 CTATTTATCCCCCTGGAATGTGCGCGCTGGCCAGAAGGAAGGGCTGGTTTTGGAAGCTCC 660

Qy       2553 CTGACCCCCGCGCTGCCCGCCCC 2575
             ||    ||| |   |   | |||
Db        661 CTTGAACCCCCCGCGGTTGGCCC 683
```

RESULT 8
BM707056
LOCUS BM707056 578 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CR1-adx-b-12-0-UI.r1 UI-E-CR1 Homo sapiens cDNA clone UI-E-CR1-adx-b-12-0-UI 5', mRNA sequence.
ACCESSION BM707056
VERSION BM707056.1 GI:19020314
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CR1-adx-b-12-0-UI"
                     /tissue_type="eye anterior segment"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CR1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CR1 is a normalized cDNA library containing the
                     following tissue(s): eye anterior segment. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AATGCCGCAT. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI)."
ORIGIN

  Query Match             17.9%;  Score 560;  DB 3;  Length 578;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-87;
  Matches  560;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1964 GGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCG 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCG 77

Qy       2024 GACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA 137

Qy       2084 AATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACA 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACA 197

Qy       2144 GCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTC 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTC 257

Qy       2204 ATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTG 2263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTG 317

Qy       2264 AACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCC 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCC 377

Qy       2324 CAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTG 437

Qy       2384 TTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTG 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTG 497

Qy       2444 TTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATC 2503
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        498 TTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTANAGCACTTAATCCTATTTATC 557

Qy       2504 CCCTGGAATGTGCGTGCTGGC 2524
             |||||||||||||||||||||
Db        558 CCCTGGAATGTGCGTGCTGGC 578


RESULT 9
W81663
LOCUS       W81663                   595 bp    mRNA    linear   EST 17-OCT-1996
DEFINITION  zd85e06.r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:347458 5', mRNA sequence.
ACCESSION   W81663
VERSION     W81663.1  GI:1392291
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 584   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1272833"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:347458"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
```

M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ORIGIN

Query Match 17.1%; Score 534.8; DB 8; Length 595;
Best Local Similarity 97.7%; Pred. No. 8.2e-83;
Matches 584; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

```
Qy       2129 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 60

Qy       2189 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 2248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 120

Qy       2249 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 180

Qy       2309 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 2368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 240

Qy       2369 TAGGGCCACCTGGTGTTTAAACA-GGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCA 2427
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        241 TAGGGCCACCTGGTGTTTAAACACGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCA 300

Qy       2428 GAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCA 2487
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCA 360

Qy       2488 CTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCA 2547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCA 420

Qy       2548 GCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGA 2607
             |||||||||||||||||||||||||||||||  |||||||||||||||| |||||||
Db        421 GCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGTAAATGTGGCATTACTGG-CCACAGA 479

Qy       2608 GGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTT 2667
             |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db        480 GGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCCTTTACTTGGGATTT 539

Qy       2668 AAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGT 2725
             ||||||| |||||| |||||||||| |||||   |||||||||||||| ||||||
Db        540 AAGAAGCCTTTAAAA-GTAATAATCCNCTGAA--GAAAATGACGTAACCNCAGCGTTT 594
```

RESULT 10
BM710194
LOCUS BM710194 500 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CQ1-agc-1-04-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone UI-E-CQ1-agc-1-04-0-UI 5', mRNA sequence.
ACCESSION BM710194
VERSION BM710194.1 GI:19023452
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CQ1-agc-1-04-0-UI"
/tissue_type="optic nerve"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CQ1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-CQ1 is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCATTAAGTG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 16.0%; Score 500; DB 3; Length 500;
Best Local Similarity 100.0%; Pred. No. 9.5e-77;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2022 CGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCC 2081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCC 60

Qy       2082 TAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTA 2141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTA 120

Qy       2142 CAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGG 2201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGG 180

Qy       2202 TCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGT 2261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGT 240

Qy       2262 TGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTC 2321
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTC 300

Qy       2322 CCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGG 2381
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGG 360

Qy       2382 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 2441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 420

Qy       2442 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 2501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 480
```

```
Qy       2502 TCCCCTGGAATGTGCGTGCT 2521
              ||||||||||||||||||||
Db        481 TCCCCTGGAATGTGCGTGCT 500


RESULT 11
AW302149/c
LOCUS       AW302149                476 bp    mRNA    linear   EST 18-JAN-2000
DEFINITION  xs01f06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2768387 3'
            similar to contains TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AW302149
VERSION     AW302149.1  GI:6711826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .476
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2768387"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
ORIGIN

  Query Match           15.1%;  Score 472.8;  DB 1;  Length 476;
  Best Local Similarity   99.6%;  Pred. No. 5.1e-72;
  Matches  474;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 417

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 357

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 297

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 237

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 177

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 117

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 57

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGC 1191
              ||||||| |||||||||||||||||||||||||||||||||||||||||| ||||
Db         56 GCGACGCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGC 1


RESULT 12
BM673230/c
LOCUS       BM673230                480 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CR1-adx-b-12-0-UI.s1 UI-E-CR1 Homo sapiens cDNA clone
            UI-E-CR1-adx-b-12-0-UI 3', mRNA sequence.
ACCESSION   BM673230
VERSION     BM673230.1  GI:18983128
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  PUBMED    8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CR1-adx-b-12-0-UI"
                     /tissue_type="eye anterior segment"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CR1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CR1 is a normalized cDNA library containing the
                     following tissue(s): eye anterior segment. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
```

Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AATGCCGCAT. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human eye anterior segment
TAG_LIB=UI-E-CR1
TAG_SEQ=AATGCCGCAT"
ORIGIN

Query Match 14.8%; Score 461.8; DB 3; Length 480;
Best Local Similarity 99.6%; Pred. No. 4.2e-70;
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2310 AAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGT 2369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGT 421

Qy       2370 AGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGA 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGA 361

Qy       2430 ACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACT 2489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACT 301

Qy       2490 TAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGC 2549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGC 241

Qy       2550 TCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGG 2609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGG 181

Qy       2610 TTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAA 2669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAA 121

Qy       2670 GAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTA 2729
             ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        120 GAAGCTTTTAAAAGGTAATAATACTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTA 61

Qy       2730 TGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||| |
Db          60 TGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCAAA 16
```

RESULT 13
BM562937
LOCUS BM562937 1376 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6566935 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5736787 5', mRNA sequence.
ACCESSION BM562937
VERSION BM562937.1 GI:18809436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1376)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12745 row: n column: 20
High quality sequence stop: 425.
FEATURES Location/Qualifiers
source 1. .1376
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5736787"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: small intestine; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Query Match 14.3%; Score 445.2; DB 3; Length 1376;
Best Local Similarity 96.9%; Pred. No. 2.7e-67;
Matches 475; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

```
Qy        453 CCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGC 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGC 60

Qy        513 ACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGG 120

Qy        573 CGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGG 180

Qy        633 GGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCC 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCC 240

Qy        693 CGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGT 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGT 300

Qy        753 TCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCGCTGAGGGGG 812
              ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db        301 TCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCAAGCTGAGGGGG 360

Qy        813 AGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        361 AGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCCAGGGAGGAG 420

Qy        873 TGCAA-GGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGG-TGC 930
              || || | |||||||||||||   ||||| | ||||||||||||||||| ||  ||||  ||
Db        421 TGTAACGCCGCTGCTGGACGGCTTCAACCACACGACTGCGTGCTACCCCCCACTGGCCGC 480

Qy        931 TGACCGTCGG 940
              ||||||||||
Db        481 TGACCGTCGG 490
```

RESULT 14
AI375213/c
LOCUS AI375213 455 bp mRNA linear EST 18-MAR-1999
DEFINITION tc10f06.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2063459 3', mRNA sequence.
ACCESSION AI375213
VERSION AI375213.1 GI:4175203
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 587   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2063459"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
ORIGIN

  Query Match            13.8%;  Score 431.4;  DB 1;  Length 455;
  Best Local Similarity  98.5%;  Pred. No. 8.2e-65;
  Matches  446;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       2323 CCAGCCACCCGGAGATGGGGG-TAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGG 2381
             |||||||||||||||||||||| ||| || || | || || ||||||||||||||||||||
Db        455 CCAGCCACCCGGAGATGGGGGTTAAAAGAAAAACCAAAGCCTTGGGGTAGGGCCACCTGG 396

Qy       2382 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 2441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 336

Qy       2442 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 2501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 276

Qy       2502 TCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCC 2561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 TCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCC 216

Qy       2562 GCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAAT 2621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 GCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAAT 156

Qy       2622 CAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAA 2681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAA 96

Qy       2682 AGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTA 2741
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 AGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTA 36

Qy       2742 TTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             |||||||||||||||||||||||||||||||||
Db          35 TTTTAATAAAGAACGCTGGGCCATGAACTCATA 3


RESULT 15
BX919953/c
LOCUS       BX919953                792 bp    mRNA    linear   EST 07-MAY-2004
DEFINITION  BX919953 Sus Scrofa library (scan) Sus scrofa cDNA clone
            scan0008d.b.15 5prim, mRNA sequence.
ACCESSION   BX919953
VERSION     BX919953.1  GI:41136772
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villeger,S.,
            Soares,M., Bonaldo,F. and Hatey,F.
  TITLE     A Pig Normalised Multi-Tissue cDNA Library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tosser-Klopp G
            Genetique Animale
            Institut National de la Recherche Agronomique
            Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
            cedex, FRANCE
            Tel: 33 (0) 5.61.28.51.14
            Fax: 33 (0) 5.61.28.53.08
            Email: tosser@toulouse.inra.fr
            Sequence cleaned of vector, adaptator and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
            Plate: 0008  row: b  column: 15.
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="scan0008d.b.15"
                     /tissue_type="mixed"
                     /dev_stage="from embryos to adults"
                     /clone_lib="Sus Scrofa library (scan)"
                     /note="tissues: adipose tissue, brain, kidney, liver,
                     muscle, ovary, testis,  heart, hypothalamus, pancreas,
                     skin, spleen, thymus, placenta, pituitary  gland, seminal
                     vesicle, small intestine, uterus, adrenals, bulbo  uretral
                     gland, cerebral trunk, epididymis, female gonad,
                     gall-bladder,  hippocampus, large intestine, male gonad,
                     melanocytes, stomach, udder"
ORIGIN

  Query Match            11.4%;  Score 357.2;  DB 5;  Length 792;
  Best Local Similarity  73.1%;  Pred. No. 6e-52;
  Matches  569;  Conservative    0;  Mismatches  165;  Indels   44;  Gaps    7;

Qy       1274 GGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGA 1333
             | | ||||  ||||| |  ||| |||   |||||||||| ||| ||||||||||||||||
Db        792 GCGGGACTCNGACGTGGAAGACNTGCACCAGCTGGAGCGGGAGATCCTTCAGGTGGGCGA 733

Qy       1334 GATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCA 1393
              |||||| |  |||||||||||||| |||||||||||||||||||||||||| ||||||||
Db        732 AATGATCCAGGACATGGAGATGAAAGTCAACGTGCCCCGCTGGACCGTGCAGGCCCGGCA 673

Qy       1394 GGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGT------CGT 1447
              || |||||||||||||||| ||  |  |||| |||   |||| ||      |||
Db        672 GGTGGCGGGCGCCGAGCTCCTCTCTAGCGCTAGCGCGGGCGTGTCCTCTGTGGGCGGCGT 613

Qy       1448 GTCCTTGCAGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTT 1507
              |||| |  |  ||| |||   |||||||  || ||||||||||| |||| | | ||
Db        612 GTCCGTAGAACAACGCACCGGGCCCTGCGACCTGAGCAAGGCCCTGGCTGCCACCATCTT 553

Qy       1508 CGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACA 1567
              | |||||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db        552 CAGCGCCGTGCTGCTGGCGGCTGTGGCCTTGGCCGTGTGCGTGGCGAAGCTGAGCTGACC 493
```

```
Qy       1568 GACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGA------------------- 1608
              ||||||||||||| || |  ||| ||||||||||||  ||
Db        492 TGCACCCGACGGCCCCCCGAAGCTACCGCTCCCTCCCTGGAGGCACCCCTCCAAGAACAA 433

Qy       1609 --------GAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCT 1660
                     |  | |||| ||| |||| ||||| | |||| |||   ||| |||  || ||
Db        432 CTGCGACTGCTAGGACTAGGGGTGGGAGTGGGATGTGGCATGTTTCAGGAGAGCTGTTCT 373

Qy       1661 AAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGAC 1720
                ||   | |  |  | |||| | || |||||| ||  ||||| |||  || ||| ||
Db        372 CGTGCC---TATACGTGTAGGTATAGGCACGTTTCAAGCACACATTTGCTAGGACAGAAC 316

Qy       1721 ACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTAC 1780
              | ||  |||||||||||||||| ||||  |||||| ||||||| ||||  ||||||||||
Db        315 ATGG-CTTCCTCTTGCTGGCCAGGGAAGAGTTAATTTTGCGCGGGCCAGCAGGGCATTAT 257

Qy       1781 CGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACCGT-CACAGTGACCCTAG 1839
              |||||| || |||||||||||||| | ||||||||||||||||| |||||||||| ||
Db        256 CGCTAATGTATGCAGGAGCTTTATCCTCTATTAATAGAAAACCGTCCACAGTGACCCGAG 197

Qy       1840 ATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGA 1899
              || ||||| |||||||  |||| | | |||||||| |||||||| ||||||  || || |
Db        196 ATTCCTCCAAGTTAATAGGTTACCGCCTGTGCTGCTGGGGCGTGTTTACACTGAATCTGG 137

Qy       1900 GTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTT-TCTGCGTGGGACAGTTT-----GAA 1953
              | | ||| | | | ||        ||   ||| || |||| ||||| ||||||     |||
Db        136 GCTTGGTCCTCTCGCCCCTGTCAGTGATCCCATTCTCTGAGTGGGGCAGTTTGGGGGGAA 77

Qy       1954 AAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTG 2011
              ||   ||||| |||| | |||||| | |||| ||  ||| ||| | | | ||||| ||
Db         76 AAAAAGGGTGAGGTGAAATGAAGTGTTGAGAAGGTGGCTATTTTGGTGTGTGTGGGTG 19
```

Search completed: May 11, 2006, 05:27:41
Job time : 11794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 01:38:14 ; Search time 530 Seconds
(without alignments)
10470.844 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	181.4	5.8	601	3	US-09-949-016-196942	Sequence 196942,
	2	181.4	5.8	109250	3	US-09-949-016-12530	Sequence 12530, A
	3	181.4	5.8	109251	3	US-09-949-016-17321	Sequence 17321, A
	4	177.2	5.7	26760	3	US-09-949-016-15894	Sequence 15894, A
c	5	176.4	5.7	601	3	US-09-949-016-40113	Sequence 40113, A
c	6	176.4	5.7	601	3	US-09-949-016-138968	Sequence 138968,
	7	176.4	5.7	30371	3	US-09-949-016-15395	Sequence 15395, A
c	8	176.4	5.7	41589	3	US-09-949-016-12867	Sequence 12867, A
c	9	176.4	5.7	41593	3	US-09-949-016-15666	Sequence 15666, A
c	10	176	5.6	601	3	US-09-949-016-40112	Sequence 40112, A
c	11	176	5.6	601	3	US-09-949-016-138967	Sequence 138967,
	12	175.6	5.6	93778	3	US-09-949-016-15096	Sequence 15096, A
c	13	174.8	5.6	23766	3	US-09-949-016-13569	Sequence 13569, A
c	14	174.8	5.6	74545	3	US-09-949-002-606	Sequence 606, App
c	15	174.8	5.6	125192	3	US-09-949-016-14120	Sequence 14120, A
c	16	174.8	5.6	131860	3	US-09-949-002-730	Sequence 730, App
c	17	174.2	5.6	98828	3	US-09-949-016-16630	Sequence 16630, A
	18	174	5.6	5371	3	US-09-949-016-16996	Sequence 16996, A
c	19	174	5.6	152481	3	US-09-949-016-12521	Sequence 12521, A
c	20	174	5.6	152798	3	US-09-949-016-12775	Sequence 12775, A
c	21	174	5.6	152822	3	US-09-949-016-17518	Sequence 17518, A
c	22	174	5.6	152822	3	US-09-949-016-17519	Sequence 17519, A
c	23	173.8	5.6	18591	3	US-09-949-016-14719	Sequence 14719, A
c	24	173.2	5.5	601	3	US-09-949-016-81759	Sequence 81759, A
c	25	173.2	5.5	48018	3	US-09-949-016-17573	Sequence 17573, A
	26	173.2	5.5	67386	3	US-09-949-016-16519	Sequence 16519, A
	27	173.2	5.5	146095	3	US-09-949-016-12872	Sequence 12872, A
	28	173.2	5.5	146104	3	US-09-949-016-13239	Sequence 13239, A
	29	173	5.5	59828	3	US-09-949-016-16238	Sequence 16238, A
c	30	172.8	5.5	36816	3	US-09-949-016-16218	Sequence 16218, A
c	31	172.4	5.5	35688	3	US-09-949-016-16873	Sequence 16873, A
c	32	172.4	5.5	49848	3	US-09-949-016-15675	Sequence 15675, A
	33	172.4	5.5	74096	3	US-09-949-016-11785	Sequence 11785, A
	34	172.4	5.5	74097	3	US-09-949-016-16239	Sequence 16239, A
c	35	172.2	5.5	18448	3	US-09-949-016-15345	Sequence 15345, A
c	36	172	5.5	60376	3	US-09-949-016-12423	Sequence 12423, A
	37	171.8	5.5	282	2	US-08-133-629-8	Sequence 8, Appli
	38	171.8	5.5	601	3	US-09-949-016-126655	Sequence 126655,
	39	171.8	5.5	35337	3	US-09-949-016-17249	Sequence 17249, A
	40	171.6	5.5	601	3	US-09-949-016-42392	Sequence 42392, A
c	41	171.6	5.5	38368	3	US-09-949-016-12958	Sequence 12958, A
c	42	171.6	5.5	58782	3	US-09-949-016-16851	Sequence 16851, A
	43	171.6	5.5	107941	3	US-09-949-016-14206	Sequence 14206, A
	44	171.6	5.5	109378	3	US-09-949-016-12391	Sequence 12391, A
c	45	171.4	5.5	20347	3	US-09-949-016-16752	Sequence 16752, A

ALIGNMENTS

RESULT 1
US-09-949-016-196942/c
; Sequence 196942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196942
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196942

Query Match 5.8%; Score 181.4; DB 3; Length 601;
Best Local Similarity 88.5%; Pred. No. 2.7e-29;
Matches 208; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

```
Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAA 2947
             |||| |||||||||||| ||||| | |||||||||||||||||||||||| ||| | ||||
Db        536 TGAGACAGGTGGATCACTTGAGGTCCGGAGTTCGAGACCAGCCTGGCCAGCATGGTGAAA 477

Qy       2948 CCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCC 3007
             ||| |||||||||||||||| ||||||| || |||  |||||||| |||||||||  |||
Db        476 CCCCATCTCTACTAAAAAATACAAAAATTAGCTGGGTGTGGTGGCACATGCCTGTAATCC 417

Qy       3008 CAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAAT 3067
             ||||||||| ||||| |||||||||||| || |||||||| ||||||| || |||||||
Db        416 CAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCAGGAGGCGGGGGTTGCAGT 357

Qy       3068 GAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             ||| |||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        356 GAGCCAAGATCATGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 302
```

RESULT 2

```
US-09-949-016-12530
; Sequence 12530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12530
;   LENGTH: 109250
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(109250)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12530

  Query Match             5.8%;  Score 181.4;  DB 3;  Length 109250;
  Best Local Similarity   88.5%;  Pred. No. 1.5e-28;
  Matches  208;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAA 2947
             |||| |||||||||||| ||||| | ||||||||||||||||||||||| ||| | ||||
Db      78099 TGAGACAGGTGGATCACTTGAGGTCCGGAGTTCGAGACCAGCCTGGCCAGCATGGTGAAA 78158

Qy       2948 CCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCC 3007
             ||| |||||||||||||||| ||||||| || |||  ||||||| |||||||||  |||
Db      78159 CCCCATCTCTACTAAAAAATACAAAAATTAGCTGGGTGTGGTGGCACATGCCTGTAATCC 78218

Qy       3008 CAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAAT 3067
             ||||||||| ||||| |||||||||||| || ||||||| ||||||| || ||||||| |
Db      78219 CAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCAGGAGGCGGGGGTTGCAGT 78278

Qy       3068 GAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              ||| |||||||  ||||||||||||||||||||||||||||||||||||||||||
Db      78279 GAGCCAAGATCATGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 78333


RESULT 3
US-09-949-016-17321
; Sequence 17321, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17321
;   LENGTH: 109251
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(109251)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17321

  Query Match             5.8%;  Score 181.4;  DB 3;  Length 109251;
  Best Local Similarity   88.5%;  Pred. No. 1.5e-28;
  Matches  208;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAA 2947
             |||| |||||||||||| ||||| | ||||||||||||||||||||||| ||| | ||||
Db      78099 TGAGACAGGTGGATCACTTGAGGTCCGGAGTTCGAGACCAGCCTGGCCAGCATGGTGAAA 78158

Qy       2948 CCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCC 3007
             ||| |||||||||||||||| ||||||| || |||  ||||||| |||||||||  |||
Db      78159 CCCCATCTCTACTAAAAAATACAAAAATTAGCTGGGTGTGGTGGCACATGCCTGTAATCC 78218

Qy       3008 CAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAAT 3067
             ||||||||| ||||| |||||||||||| || ||||||| ||||||| || ||||||| |
Db      78219 CAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCAGGAGGCGGGGGTTGCAGT 78278

Qy       3068 GAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              ||| |||||||  ||||||||||||||||||||||||||||||||||||||||||
Db      78279 GAGCCAAGATCATGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 78333


RESULT 4
US-09-949-016-15894
; Sequence 15894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15894
;   LENGTH: 26760
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15894

  Query Match             5.7%;  Score 177.2;  DB 3;  Length 26760;
  Best Local Similarity   87.6%;  Pred. No. 7.4e-28;
  Matches  205;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAAC 2948
             |||||||||||||||||||| |||||||| |||||||||||||||||| || ||||
Db      21533 GAGGCAGGTGGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGCAAAAC 21592

Qy       2949 CCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCC 3008
             || |||||||||||||    ||||||| ||||||| ||||||| ||||||| | ||||
Db      21593 CCCATCTCTACTAAAAATACAAAAAATTAGCCGGGCGTGGTGGCACATGCCTATAATCCC 21652

Qy       3009 AGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATG 3068
             ||||||| |||||| ||||||| |||| || ||||||    |||||||||||||||| ||
Db      21653 AGCTACTTGGGAGGCTGAGGCACGAGAATCACTTGAACTTGGGAGGTGGAGGTTGCAGTG 21712

Qy       3069 AG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              || | |||||||||||||||||||| |||||||||||||| ||||||||||
Db      21713 AGCCGAGATTGTGCCACTGCACTCCAGTCTGGGCAACAGAGCGAGACTCTGTCT 21766


RESULT 5
```

```
US-09-949-016-40113/c
; Sequence 40113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40113
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-40113

  Query Match             5.7%;  Score 176.4;  DB 3;  Length 601;
  Best Local Similarity   83.5%;  Pred. No. 3.1e-28;
  Matches  212;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| ||  |||     |  ||||| |||||||||| |||||||||||||||| ||||||||
Db        259 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 200

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||||||||||||||  | | ||||||||||||    ||||||| ||| ||||||||
Db        199 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 140

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             ||||||  |||||  |||||||||||||||||| ||||||| ||| ||||||||  ||
Db        139 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 80

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             |||||  |||||||| |||| ||||||   |||||||||||||||||||||||||| ||
Db          79 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACACAG 20

Qy       3109 CAAGACTCTGTCTC 3122
             |||||||| |||||
Db          19 CAAGACTCCGTCTC 6


RESULT 6
US-09-949-016-138968/c
; Sequence 138968, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 138968
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-138968

  Query Match             5.7%;  Score 176.4;  DB 3;  Length 601;
  Best Local Similarity   83.5%;  Pred. No. 3.1e-28;
  Matches  212;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| ||  |||     |  ||||| |||||||||| |||||||||||||||| ||||||||
Db        259 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 200

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||||||||||||||  | | ||||||||||||    ||||||| ||| ||||||||
Db        199 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 140

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             ||||||  |||||  |||||||||||||||||| ||||||| ||| ||||||||  ||
Db        139 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 80

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             |||||  |||||||| |||| ||||||   |||||||||||||||||||||||||| ||
Db          79 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACACAG 20

Qy       3109 CAAGACTCTGTCTC 3122
             |||||||| |||||
Db          19 CAAGACTCCGTCTC 6


RESULT 7
US-09-949-016-15395
; Sequence 15395, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15395
;   LENGTH: 30371
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15395

  Query Match             5.7%;  Score 176.4;  DB 3;  Length 30371;
  Best Local Similarity   83.5%;  Pred. No. 1.1e-27;
  Matches  212;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| |  |||      |||||| ||  |||||||||||| |||||||| |||| ||||
Db      26920 CCCAGCACTTTGGGAGGCTGAGGCGGGCAGATCACCTGAGGTCAGGAGTTTGAGATCAGC 26979

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||| ||| ||||||||| |||||||||||||||| ||||||| ||||||||||||
Db      26980 CTGGCCAGCATGGCGAAACCCCATCTCTACTAAAAAATACAAAAATTAGCCGGGCATGGT 27039

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             ||| |  ||||||  |||||||||||| |||| | ||||||||||||  | ||||||  |
Db      27040 GGCACGGGCCTGTAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAACCACTTGAACCCG 27099

Qy       3050 GGAGGTGGAGGTTGCAATGAGC-AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             |||||||||||||||| |||||  |||||||||| |||||||||||| ||||||||||
Db      27100 GGAGGTGGAGGTTGCAGTGAGCTGAGATTGTGCCATTGCACTCCAGCCCGGGCAACAGAG 27159

Qy       3109 CAAGACTCTGTCTC 3122
             ||| ||||||||||
```

```
Db      27160 CAAAACTCTGTCTC 27173


RESULT 8
US-09-949-016-12867/c
; Sequence 12867, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12867
;   LENGTH: 41589
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(41589)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12867

  Query Match             5.7%;  Score 176.4;  DB 3;  Length 41589;
  Best Local Similarity   83.5%;  Pred. No. 1.3e-27;
  Matches  212;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| ||  |||     |  ||||| |||||||||| |||||||||||||| ||||||||
Db       8267 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 8208

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             |||||||||||||||||  | | ||||||||||||    ||||||| ||| ||||||||
Db       8207 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 8148

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             ||||||  |||||  |||||||||||||||||| |||||||| |||  |||||||  ||
Db       8147 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 8088

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             |||||  |||||||| |||| ||||||   |||||||||||||||||||||||||||| ||
Db       8087 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACACAG 8028

Qy       3109 CAAGACTCTGTCTC 3122
             |||||||| |||||
Db       8027 CAAGACTCCGTCTC 8014


RESULT 9
US-09-949-016-15666/c
; Sequence 15666, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15666
;   LENGTH: 41593
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(41593)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15666

  Query Match             5.7%;  Score 176.4;  DB 3;  Length 41593;
  Best Local Similarity   83.5%;  Pred. No. 1.3e-27;
  Matches  212;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| ||  |||     |  ||||| |||||||||| |||||||||||||| ||||||||
Db       8267 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 8208

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             |||||||||||||||||  | | ||||||||||||    ||||||| ||| ||||||||
Db       8207 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 8148

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             ||||||  |||||  |||||||||||||||||| |||||||| |||  |||||||  ||
Db       8147 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 8088

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             |||||  |||||||| |||| ||||||   |||||||||||||||||||||||||||| ||
Db       8087 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCAACACAG 8028

Qy       3109 CAAGACTCTGTCTC 3122
             |||||||| |||||
Db       8027 CAAGACTCCGTCTC 8014


RESULT 10
US-09-949-016-40112/c
; Sequence 40112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40112
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-40112

  Query Match             5.6%;  Score 176;  DB 3;  Length 601;
  Best Local Similarity   83.1%;  Pred. No. 3.8e-28;
  Matches  211;  Conservative    1;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             ||| ||  |||     |  ||||| |||||||||| |||||||||||||| ||||||||
Db        512 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 453

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             |||||||||||||||||  | | ||||||||||||    ||||||| ||| ||||||||
```

```
Db        452 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 393

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              ||||||  |||||  |||||||||||||||||| ||||||||  |||  ||||||||  ||
Db        392 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 333

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              |||||  |||||||||  ||||  ||||||   :|||||||||||||||||||||||||| ||
Db        332 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACRCCACTGCACTCCAGCCTGGGCAACACAG 273

Qy       3109 CAAGACTCTGTCTC 3122
              ||||||||  |||||
Db        272 CAAGACTCCGTCTC 259


RESULT 11
US-09-949-016-138967/c
; Sequence 138967, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 138967
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-138967

  Query Match             5.6%;  Score 176;  DB 3;  Length 601;
  Best Local Similarity   83.1%;  Pred. No. 3.8e-28;
  Matches  211;  Conservative    1;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
              |||  ||  |||     |  ||||| |||||||||| ||||||||||||||| ||||||||
Db        512 CCCAACACTTTGGGAATCCGAGGCGGGTGGATCACTTGAGGCCAGGAGTTCAAGACCAGC 453

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
              ||||||||||||||||||  | | ||||||||||||     |||||||  ||| ||||||||
Db        452 CTGGCCAACATAGCGAAATTCCACCTCTACTAAAAATACAAAAAATTAGCCAGGCATGGT 393

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              ||||||  |||||  |||||||||||||||||| ||||||||  |||  ||||||||  ||
Db        392 GGCGCACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGAAGAACCGCTTGAACCCA 333

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              |||||  |||||||||  ||||  ||||||   :|||||||||||||||||||||||||| ||
Db        332 GGAGGCAGAGGTTGCAGTGAGCCAAGATCACRCCACTGCACTCCAGCCTGGGCAACACAG 273

Qy       3109 CAAGACTCTGTCTC 3122
              ||||||||  |||||
Db        272 CAAGACTCCGTCTC 259


RESULT 12
US-09-949-016-15096
; Sequence 15096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15096
;   LENGTH: 93778
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(93778)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15096

  Query Match             5.6%;  Score 175.6;  DB 3;  Length 93778;
  Best Local Similarity   87.2%;  Pred. No. 2.4e-27;
  Matches  204;  Conservative    0;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACC 2949
              |||||||||||||||  ||||| |||||||||||||||||||||||||| | | || |||
Db      74283 AGGCAGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACCTGGTGATACC 74342

Qy       2950 CGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCA 3009
              | |||||||||| |||| | ||||||| ||| ||||| |||| || ||||||| ||||||
Db      74343 CCATCTCTACTAGAAAAAAAAAAAATTAGCCAGGCATAGTGGTGCGTGCCTGTAGTCCCA 74402

Qy       3010 GCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATGA 3069
              ||||||||||||| |||||||||||| || ||||||  ||||||| |||| ||||| |||
Db      74403 GCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACACAGGAGGCGGAGATTGCAGTGA 74462

Qy       3070 GC-AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||  |||| | ||||||||||||||||||||  ||||||||||||||||||||
Db      74463 GCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCTGTCTC 74516


RESULT 13
US-09-949-016-13569/c
; Sequence 13569, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13569
;   LENGTH: 23766
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13569

  Query Match             5.6%;  Score 174.8;  DB 3;  Length 23766;
  Best Local Similarity   83.1%;  Pred. No. 2.3e-27;
  Matches  211;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
              |||  | ||||       ||||  || |||||||||||||| |||||||| || ||||||
```

```
Db       4585 CCCAGCATTTTGGGGGGCCGAGGTGGGGGGATCACCTGAGGTCAGGAGTTTGAAACCAGC 4526

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
              |||| |||  | || |||| | |||||||||||||||||| |||||||  ||||| |||||
Db       4525 CTGGTCAATGTGGCAAAACACCATCTCTACTAAAAAATACAAAAATTATCCGGGTATGGT 4466

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              |||   |||||| |||||||||||| |||||| |||||||||||| || ||||||
Db       4465 GGCATGTGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCTG 4406

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              |||||||||||||||| |||| |||||||||||| |||||||||||||||||||||||
Db       4405 GGAGGTGGAGGTTGCAGTGAGTCAAGATTGTGCCATTGCACTCCAGCCTGGGCAACAGAG 4346

Qy       3109 CAAGACTCTGTCTC 3122
              ||||||||||||||
Db       4345 CAAGACTCTGTCTC 4332


RESULT 14
US-09-949-002-606/c
; Sequence 606, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 606
;   LENGTH: 74545
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-606

  Query Match             5.6%;  Score 174.8;  DB 3;  Length 74545;
  Best Local Similarity   83.1%;  Pred. No. 3.3e-27;
  Matches  211;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
              |||  ||  |||       ||| |  || |||||||||||| |||||||| ||||||||
Db      11479 CCCAACAGTTTAAAAGGCTGAAGTGGGAGGATCACCTGAGGTCAGGAGTTCAAGACCAGC 11420

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
              |||||||||||| ||||||| |||||||||||||| | ||||||| ||| ||   ||||
Db      11419 CTGGCCAACATAGAGAAACCCCATCTCTACTAAAAAACACAAAAATTAGCCAGGAGTGGT 11360

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              | |    ||||||  |||||||||||||||||| |||||||||| |||||||||  |
Db      11359 GACAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCG 11300

Qy       3050 GGAGGTGGAGGTTGCAATGAGC-AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              ||||| |||||||||| |||||  |||||| |||||||||||||||||||| ||||||
Db      11299 GGAGGCGGAGGTTGCAGTGAGCTGAGATTGCGCCACTGCACTCCAGCCTGGGCGACAGAG 11240

Qy       3109 CAAGACTCTGTCTC 3122
              |||||||| |||||
Db      11239 CAAGACTCCGTCTC 11226


RESULT 15
US-09-949-016-14120/c
; Sequence 14120, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14120
;   LENGTH: 125192
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(125192)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14120

  Query Match             5.6%;  Score 174.8;  DB 3;  Length 125192;
  Best Local Similarity   83.1%;  Pred. No. 4e-27;
  Matches  211;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
              |||  |  ||||       ||||||||| |||||||||||||| |||||||| |||||||
Db      67067 CCCAGCACTTTTGGAGGCTGAGGCAGGCGGATCACCTGAGGTCAGGAGTTCAAGACCAGG 67008

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
              |||||||||||  | || ||||  |||||| ||||||     ||||||| ||||||| ||||
Db      67007 CTGGCCAACATGGTGAGACCCTGTCTCTATTAAAAATACAAAAAATTAGCCGGGCGTGGT 66948

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              || |||||||| |  |||||||||||||| ||| |||||||||||| || ||||||  ||
Db      66947 GGTGCATGCCTATAATCCCAGCTACTCGGAAGGGTGAGGCAGGAGAATCACTTGAACCCA 66888

Qy       3050 GGAGGTGGAGGTTGCAATGAGC-AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              ||||| |||||||||| |||||  |||| ||||||||||||||||||| ||||||||||
Db      66887 GGAGGCGGAGGTTGCAGTGAGCTGAGATCGTGCCACTGCACTCCAGCCAGGGCAACAGAG 66828

Qy       3109 CAAGACTCTGTCTC 3122
              ||||||||||||||
Db      66827 TAAGACTCTGTCTC 66814


Search completed: May 11, 2006, 06:02:58
Job time : 532 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 02:04:15 ; Search time 2407 Seconds
(without alignments)
10725.799 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	846.8	27.1	850	9	US-10-450-763-29532	Sequence 29532, A
	2	726.6	23.3	1061	8	US-10-363-345A-27371	Sequence 27371, A
c	3	726.6	23.3	1061	8	US-10-363-345A-27372	Sequence 27372, A
	4	726.6	23.3	1061	9	US-10-363-483A-27371	Sequence 27371, A
c	5	726.6	23.3	1061	9	US-10-363-483A-27372	Sequence 27372, A
c	6	706.4	22.6	708	6	US-10-029-386-20699	Sequence 20699, A
c	7	594.4	19.0	1061	8	US-10-363-345A-27369	Sequence 27369, A
	8	594.4	19.0	1061	8	US-10-363-345A-27370	Sequence 27370, A
c	9	594.4	19.0	1061	9	US-10-363-483A-27369	Sequence 27369, A
	10	594.4	19.0	1061	9	US-10-363-483A-27370	Sequence 27370, A
c	11	559.6	17.9	560	4	US-09-925-065A-740956	Sequence 740956,
c	12	556.6	17.8	557	4	US-09-925-065A-769508	Sequence 769508,
c	13	555.6	17.8	556	4	US-09-925-065A-736351	Sequence 736351,
	14	554.6	17.8	555	4	US-09-925-065A-737120	Sequence 737120,
c	15	524	16.8	524	6	US-10-029-386-6988	Sequence 6988, Ap
c	16	318.6	10.2	1349	9	US-10-450-763-8312	Sequence 8312, Ap
	17	256.2	8.2	432	9	US-10-450-763-29529	Sequence 29529, A
c	18	180.2	5.8	2306	4	US-09-925-065A-706227	Sequence 706227,
c	19	180.2	5.8	3347	5	US-10-027-632-116266	Sequence 116266,
c	20	180.2	5.8	3347	6	US-10-027-632-116266	Sequence 116266,
	21	178.6	5.7	414	4	US-09-925-065A-132933	Sequence 132933,
c	22	178.6	5.7	183334	8	US-10-741-600-17646	Sequence 17646, A
	23	177.6	5.7	637	5	US-10-027-632-221302	Sequence 221302,
	24	177.6	5.7	637	5	US-10-027-632-221303	Sequence 221303,
	25	177.6	5.7	637	5	US-10-027-632-221304	Sequence 221304,
	26	177.6	5.7	637	6	US-10-027-632-221302	Sequence 221302,
	27	177.6	5.7	637	6	US-10-027-632-221303	Sequence 221303,
	28	177.6	5.7	637	6	US-10-027-632-221304	Sequence 221304,
c	29	177.6	5.7	845	5	US-10-027-632-169470	Sequence 169470,
c	30	177.6	5.7	845	6	US-10-027-632-169470	Sequence 169470,
	31	177.2	5.7	24295	7	US-10-317-277A-4	Sequence 4, Appli
	32	176.8	5.7	721	5	US-10-027-632-11218	Sequence 11218, A
	33	176.8	5.7	721	6	US-10-027-632-11218	Sequence 11218, A
c	34	176.6	5.7	166536	9	US-10-981-277-35	Sequence 35, Appl
	35	176.4	5.7	38753	7	US-10-741-601-5767	Sequence 5767, Ap
c	36	176.4	5.7	158001	7	US-10-211-179-11	GENERAL INFORMATI
c	37	176	5.6	433	5	US-10-027-632-256213	Sequence 256213,
c	38	176	5.6	433	5	US-10-027-632-256214	Sequence 256214,
c	39	176	5.6	433	6	US-10-027-632-256213	Sequence 256213,
c	40	176	5.6	433	6	US-10-027-632-256214	Sequence 256214,
	41	176	5.6	19300	7	US-10-317-271A-4	Sequence 4, Appli
c	42	175.6	5.6	755	5	US-10-027-632-27592	Sequence 27592, A
c	43	175.6	5.6	755	5	US-10-027-632-27593	Sequence 27593, A
c	44	175.6	5.6	755	6	US-10-027-632-27592	Sequence 27592, A
c	45	175.6	5.6	755	6	US-10-027-632-27593	Sequence 27593, A

ALIGNMENTS

RESULT 1
US-10-450-763-29532
; Sequence 29532, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 29532
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (241)..(699)
; OTHER INFORMATION: 30% homologous to Leishmania major L8453.1,accession number
; OTHER INFORMATION: AC008054,Smith-Waterman Score=81.
US-10-450-763-29532

Query Match 27.1%; Score 846.8; DB 9; Length 850;
Best Local Similarity 99.8%; Pred. No. 4.9e-205;
Matches 848; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 60

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 120

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 180

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
```

```
Db        181 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 240

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 300

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 360

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 420

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGT 1195
              ||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        421 GCGACGCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGT 480

Qy       1196 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 540

Qy       1256 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 600

Qy       1316 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 660

Qy       1376 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 720

Qy       1436 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 780

Qy       1496 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 840

Qy       1556 GCTGAGCTGA 1565
              ||||||||||
Db        841 GCTGAGCTGA 850
```

```
RESULT 2
US-10-363-345A-27371
; Sequence 27371, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27371
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27371
US-10-363-345A-27371

  Query Match             23.3%;  Score 726.6;  DB 8;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 2.1e-174;
```

```
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554
              | |||   |||||||| |   |||||||||||||| || |   | || |||| ||||||
Db          1 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 60

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614
              | ||||||||  |||||||| |||| ||||||||||| |||||||||   ||||||
Db         61 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 120

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674
              |||||||||||||||||||| || |||||||| |   | ||||||   ||| || |||
Db        121 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 180

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734
              |||| ||||||||||| |||||||| | ||||||||||||||||||| |||||||||
Db        181 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 240

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794
              |||||  |||||||||||||||| |||||||    |||| ||||||||||||||  | ||
Db        241 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 300

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854
              || |||| |||||||||| |||||||||||||| ||||   || |||||  | |
Db        301 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 360

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914
              |||||||||||||||||| |||||| || |||||||| | || ||||||| ||||||
Db        361 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 420

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974
              ||  |  |  ||||| ||| |||||||| ||| |||| ||||| |||||
Db        421 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 480

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034
              || |||||||||  |||||||| ||||| |||||||||  |  ||||  ||| ||| |
Db        481 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 540

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094
              | |||| ||||||| ||||| ||| || |||||||||| |||||||||||| | |||
Db        541 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 600

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154
              ||||  || ||||| ||  |||| || |||||||||| |||||||| || ||||| ||
Db        601 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 660

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214
              ||||  ||||||  || || |||||  |||||||  || |||||||  | ||||||| ||||
Db        661 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 720

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274
              |  |  | |||||||||||||||||| |||| |  || ||  ||||| ||||||||||
Db        721 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 780

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334
              ||||| ||||||||||||||||  |||||||| ||||||||||||  || ||||||||||
Db        781 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 840

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394
              |||||||| || |||||||||||||| ||||||   || |||| |||| |||  ||| ||
Db        841 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 900

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454
              |||||||||| |||| |  |||  ||||| |||| |||    |  |||||||||  |||
Db        901 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 960

Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514
              |||||||||||||||||||||||     |||||||    ||| ||  ||  ||||||||| |
Db        961 TAGGAGCGCGGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 1020

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||| || |||||| |||||    ||| ||||||||||||||
Db       1021 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1061
```

```
RESULT 3
US-10-363-345A-27372/c
; Sequence 27372, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27372
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27372
US-10-363-345A-27372

  Query Match             23.3%;  Score 726.6;  DB 8;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 2.1e-174;
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554
              | |||   |||||||||| |   |||||||||||||| || |   | || |||| ||||||
Db       1061 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 1002

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614
                | ||||||||  |||||||| ||||| |||||||||||| ||||||||||  ||||||
Db       1001 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 942

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674
              |||||||||||||||||||||| || |||||||| |   | ||||||   ||| || |||
Db        941 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 882

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734
              |||| ||||||||||   |||||||| | ||||||||||||||||||||| |||||||||
Db        881 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 822

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794
              |||||  |||||||||||||||| |||||||    |||| ||||||||||||||||  | ||
Db        821 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 762

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854
                || |||| |||||||||| |||||||||||||||| ||||   || |||||   | |
Db        761 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 702

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914
              |||||||||||||||||||| |||||| || |||||||| | || ||||||| ||||||
Db        701 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 642

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974
              ||   |  |  ||||| ||| |||||||| |||||||| ||| ||||  ||||| |||||
Db        641 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 582

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034
               || |||||||||  |||||||| |||| |||||||| |  ||||  ||| ||| | |
Db        581 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 522

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094
              | |||| ||||||| |||||  ||| || |||||||||| |||||||||||||| | |||
Db        521 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 462

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154
              ||||  || |||||  ||  ||||  || |||||||||| |||||||| || ||||| ||
Db        461 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 402

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214
              |||| ||||||   || || ||||| |||||||  || |||||||  | |||||||| ||||
Db        401 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 342

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274
              |  |  | |||||||||||||||||| |||| |  || ||  ||||| ||||||||||
Db        341 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 282

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334
              ||||| ||||||||||||||  |||||||| ||||||||||||||  || ||||||||||
Db        281 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 222

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394
              |||||||| || |||||||||||||| ||||||   || |||| |||| |||  ||| ||
Db        221 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 162

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454
              |||||||||| |||| |  |||  ||||| |||| |||    |  |||||||||  |||
Db        161 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 102

Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514
              |||||||||||||||||||||||     |||||||   ||| ||  ||  |||||||| |
Db        101 TAGGAGCGCGGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 42

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||| || |||||| |||||   ||| ||||||||||||||||
Db          41 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1

RESULT 4
US-10-363-483A-27371
; Sequence 27371, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27371
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27371
US-10-363-483A-27371

  Query Match             23.3%;  Score 726.6;  DB 9;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 2.1e-174;
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554
              | |||   |||||||||| |   |||||||||||||| || |   | || |||| ||||||
Db          1 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 60

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614
                | ||||||||  |||||||| ||||| |||||||||||  ||||||  ||||||
Db         61 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 120

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674
              |||||||||||||||||||||| || |||||||| |   | ||||||   ||| || |||
Db        121 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 180

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734
              |||| ||||||||||   |||||||| | ||||||||||||||||||||| |||||||||
Db        181 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 240

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794
```

```
Db        241 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 300

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854

Db        301 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 360

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914

Db        361 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 420

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974

Db        421 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 480

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034

Db        481 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 540

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094

Db        541 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 600

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154

Db        601 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 660

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214

Db        661 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 720

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274

Db        721 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 780

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334

Db        781 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 840

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394

Db        841 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 900

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454

Db        901 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 960

Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514

Db        961 TAGGAGCGCGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 1020

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555

Db       1021 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1061
```

```
RESULT 5
US-10-363-483A-27372/c
; Sequence 27372, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27372
;   LENGTH: 1061
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27372
US-10-363-483A-27372

  Query Match           23.3%;  Score 726.6;  DB 9;  Length 1061;
  Best Local Similarity   80.3%;  Pred. No. 2.1e-174;
  Matches  852;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy        495 GCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAG 554

Db       1061 GTTTGTTTCGGAGTTGGTATTTACGGAGGATGGGGATCGTATTTTTAGTTTCGTAGGGAG 1002

Qy        555 CCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGA 614

Db       1001 TTATCGTGGAGGTTAGGGCGGTGTAGAGATACGACGTGTGATTCGGAGTGCGTTTGGGGA 942

Qy        615 GGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGA 674

Db        941 GGATGGACGAGGGAGCGGGGGATCGTTAACGGGGTTTTTTTTGCGCGTTTCGTTCGTAGA 882

Qy        675 GGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTC 734

Db        881 GGCGTACGTCGAGGGTTTCGGGCGGGTTTCGTGGACGTTGGCGGTAGCGTCGAGCGAGTT 822

Qy        735 ACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACAT 794

Db        821 ACGGATTATGAAGAGCGTTCGTGTCGCGCGGTTTAAGGTCGGGATGGGGGTTAGTTATAT 762

Qy        795 CCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACC 854

Db        761 TTTGTCGCGTTGAGGGGGAGGTTAACGGGCGCGGGCGGTCGGGTTTAGTCGGAGTTTATC 702

Qy        855 GCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGC 914

Db        701 GCGATGGCGAGGGAGGAGTGTAAGGCGTTGTTGGACGGGTTTAATAAGACGATTGCGTGT 642

Qy        915 TACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAG 974

Db        641 TATTATTATTTGGTGTTGATCGTCGGTGGTTCGGCGGATTCGTAGAATTTGCGGTAGGAG 582

Qy        975 CTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACT 1034

Db        581 TTGTAAAAGACGCGTTAGAAGGCGTAGGAGTTGGCGGTGTTTATTTGCGTTCGGTTGATT 522

Qy       1035 GCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGG 1094

Db        521 GTTGTGTTGCGCGATCGGGGTTTGGTCGTCGACGAGCGCGTCGAGTTCGAGCGGTTTTGG 462

Qy       1095 GTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTG 1154

Db        461 GTGGTTTTTTCGGGTTGTTTGGATTTGTTGGAAGCGGATATGCGACGCGCGTTGGAGTTG 402

Qy       1155 GGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGC 1214

Db        401 GGCGTCGCGTTTTCGTTGTACGCGTCGCGGCGGTCGTTGGTGCGTATAGGTGTGGTTGGC 342

Qy       1215 GCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAG 1274

Db        341 GTTTTTTTCGGCGTGGCGGCGCGCGCGTTGAGTATTCGTAGTTTGCGGTTCGAGGCGGAG 282

Qy       1275 GGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334

Db        281 GGCGATTTCGACGTCGCGGATTTGCGGGAGTTGGAGCGCGAGGTTTTTTAGGTGGGCGAG 222

Qy       1335 ATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAG 1394

Db        221 ATGATCGATAATATGGAGATGAAGGTTAACGTGTTTCGTTGGATCGTGTAAGTTCGGTAG 162

Qy       1395 GCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTG 1454

Db        161 GCGGCGGGCGTCGAGTTTTTGTTTACGGTTAGCGTCGGTTTTTTTTCGGTCGTGTTTTTG 102
```

```
Qy       1455 CAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCC 1514
             ||||||||||||||||||||||    |||||||   ||| ||  ||  |||||||| |
Db        101 TAGGAGCGCGGGGGGGGGTTGCGATTTTAGGAAGGTTTTGGTCGTTATTTTTTTCGGCGTC 42

Qy       1515 GTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
             ||| || |||||| |||||   ||| ||||||||||||||
Db         41 GTGTTGTTGGCGGTTGTGGTTTTAGTCGTGTGCGTGGCGAA 1
```

RESULT 6
US-10-029-386-20699/c
; Sequence 20699, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20699
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010615.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EST_HUMAN HIT: AW302149.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P31735, EVALUE 5.50e+00
US-10-029-386-20699

Query Match 22.6%; Score 706.4; DB 6; Length 708;
Best Local Similarity 99.9%; Pred. No. 2.6e-169;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 649

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 589

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 529

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 469

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 409

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        408 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 349

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 289
```

```
Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 229

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 169

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 109

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 49

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1
```

RESULT 7
US-10-363-345A-27369/c
; Sequence 27369, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27369
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 27369
US-10-363-345A-27369

Query Match 19.0%; Score 594.4; DB 8; Length 1061;
Best Local Similarity 72.5%; Pred. No. 9.8e-141;
Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

```
Qy        496 CTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGC 555
              ||| ||||| | ||  |||||||| |  ||    |||||||||||| |||||||   | |
Db       1060 CTTACCCCGAAATTAACACCCACGAAAAATAAAAACCGCACCCTCAACTTCGCAAAAAAC 1001

Qy        556 CACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAG 615
              |||||  |  |||   || | || | ||||||||| | ||||| | | |||||    |
Db       1000 CACCGTAAAAACCAAAACGATACAAAAACACGACGTATAACTCGAAATACGCCTAAAAAA 941

Qy        616 GATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAG 675
              ||  ||||   | ||    |||||||||||   |||||||| ||||||||||||||| |
Db        940 AATAAACGAAAAAACGAAAAACCGCTAACGAAACTCCCTCTACGCGCCCCGTCCGCAAAA 881

Qy        676 GCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCA 735
              ||||||||||   |||||  ||  ||||||  |||||  || || ||||||  ||| |||
Db        880 ACGCACGTCGAAAATCCCGAACGAACTCCGTAAACGTTAACGATAACGCCGAACGAATCA 821

Qy        736 CGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATC 795
              || ||||| || | ||||||| ||||||| |||||  |||  ||      ||| |||||||
Db        820 CGAACCATAAAAAAACGTTCGTACCGCGCGACCCAAAACCGAAATAAAAATTAACCACATC 761

Qy        796 CTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCG 855
              || ||||||| |     |  ||||||  ||||  || |||  |||| ||| | |||||||
Db        760 CTACCGCGCTAAAAAAAAAACTAACGAACGCGAACGACCGAACCCAACCGAAACCCACCG 701

Qy        856 CGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCT 915
```

```
Db        700 CGATAACGAAAAAAAAATACAAAACGCTACTAAACGAACTCAACAAAACGACTACGTACT 641

Qy        916 ACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGC 975

Db        640 ACCACCACCTAATACTAACCGTCGATAACTCGACGAACTCGCAAAACCTACGACAAAAAC 581

Qy        976 TGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTG 1035

Db        580 TACAAAAAACGCGCCAAAAAACGCAAAAACTAACGATATCCACCTACGCCCGACTAACTA 521

Qy       1036 CTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGG 1095

Db        520 CTATACTACGCGACCGAAACCTAACCGCCGACGAACGCGCCGAATTCGAACGACTCTAAA 461

Qy       1096 TGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGG 1155

Db        460 TAACCTTCTCGAACTACCTAAACCTACTAAAAACGAACATACGACGCGCGCTAAAACTAA 401

Qy       1156 GCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCG 1215

Db        400 ACGCCGCGTTCCCGCTACACGCGCCGCGACGACCGCTAATACGCACAAATATAACTAACG 341

Qy       1216 CCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGG 1275

Db        340 CCTCCTCCGACGTAACGACGCGCGCGCTAAACACCCGCAACCTACGACTCGAAACGAAAA 281

Qy       1276 GCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGA 1335

Db        280 ACGACTTCGACGTCGCGAACCTACGAAAACTAAAACGCGAAATCCTTCAAATAAACGAAA 221

Qy       1336 TGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGG 1395

Db        220 TAATCGACAACATAAAAATAAAAATCAACGTACCCCGCTAAACCGTACAAACCCGACAAA 161

Qy       1396 CGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGC 1455

Db        160 CGACGAACGCCGAACTCCTATCCACGATCAACGCCGACCCCTCCTCGATCGTATCCTTAC 101

Qy       1456 AGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCG 1515

Db         100 AAAAACGCGAAAAAAATTACGACCCCAAAAAAACCCTAACCGCCATCCTTTTCGACGCCG 41

Qy       1516 TGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555

Db          40 TACTACTAACGACTATAACCCTAACCGTATACGTAACGAA 1


RESULT 8
US-10-363-345A-27370
; Sequence 27370, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27370
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27370
US-10-363-345A-27370

  Query Match             19.0%;  Score 594.4;  DB 8;  Length 1061;
  Best Local Similarity   72.5%;  Pred. No. 9.8e-141;

  Matches  769;  Conservative    0;  Mismatches  291;  Indels    0;  Gaps    0;

Qy        496 CTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGC 555

Db          2 CTTACCCCGAAATTAACACCCACGAAAAATAAAAACCGCACCCTCAACTTCGCAAAAAAC 61

Qy        556 CACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAG 615

Db         62 CACCGTAAAAACCAAAACGATACAAAAACACGACGTATAACTCGAAATACGCCTAAAAAA 121

Qy        616 GATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAG 675

Db        122 AATAAACGAAAAAACGAAAAACCGCTAACGAAACTCCCTCTACGCGCCCCGTCCGCAAAA 181

Qy        676 GCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCA 735

Db        182 ACGCACGTCGAAAATCCCGAACGAACTCCGTAAACGTTAACGATAACGCCGAACGAATCA 241

Qy        736 CGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATC 795

Db        242 CGAACCATAAAAAACGTTCGTACCGCGCGACCCAAAACCGAAATAAAAATTAACCACATC 301

Qy        796 CTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCG 855

Db        302 CTACCGCGCTAAAAAAAAAACTAACGAACGCGAACGACCGAACCCAACCGAAACCCACCG 361

Qy        856 CGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCT 915

Db        362 CGATAACGAAAAAAAAATACAAAACGCTACTAAACGAACTCAACAAAACGACTACGTACT 421

Qy        916 ACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGC 975

Db        422 ACCACCACCTAATACTAACCGTCGATAACTCGACGAACTCGCAAAACCTACGACAAAAAC 481

Qy        976 TGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTG 1035

Db        482 TACAAAAAACGCGCCAAAAAACGCAAAAACTAACGATATCCACCTACGCCCGACTAACTA 541

Qy       1036 CTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGG 1095

Db        542 CTATACTACGCGACCGAAACCTAACCGCCGACGAACGCGCCGAATTCGAACGACTCTAAA 601

Qy       1096 TGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGG 1155

Db        602 TAACCTTCTCGAACTACCTAAACCTACTAAAAACGAACATACGACGCGCGCTAAAACTAA 661

Qy       1156 GCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCG 1215

Db        662 ACGCCGCGTTCCCGCTACACGCGCCGCGACGACCGCTAATACGCACAAATATAACTAACG 721

Qy       1216 CCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGG 1275

Db        722 CCTCCTCCGACGTAACGACGCGCGCGCTAAACACCCGCAACCTACGACTCGAAACGAAAA 781

Qy       1276 GCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGA 1335

Db        782 ACGACTTCGACGTCGCGAACCTACGAAAACTAAAACGCGAAATCCTTCAAATAAACGAAA 841

Qy       1336 TGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGG 1395

Db        842 TAATCGACAACATAAAAATAAAAATCAACGTACCCCGCTAAACCGTACAAACCCGACAAA 901

Qy       1396 CGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGC 1455

Db        902 CGACGAACGCCGAACTCCTATCCACGATCAACGCCGACCCCTCCTCGATCGTATCCTTAC 961

Qy       1456 AGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCG 1515

Db        962 AAAAACGCGAAAAAAATTACGACCCCAAAAAAACCCTAACCGCCATCCTTTTCGACGCCG 1021

Qy       1516 TGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555

Db       1022 TACTACTAACGACTATAACCCTAACCGTATACGTAACGAA 1061
```

```
RESULT 9
US-10-363-483A-27369/c
; Sequence 27369, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27369
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27369
US-10-363-483A-27369

  Query Match            19.0%;  Score 594.4;  DB 9;  Length 1061;
  Best Local Similarity   72.5%;  Pred. No. 9.8e-141;
  Matches  769;  Conservative    0;  Mismatches  291;  Indels    0;  Gaps    0;

Qy        496 CTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGC 555
              ||| ||||| | ||  |||||||| |  ||    |||||||||||| |||||||   | |
Db       1060 CTTACCCCGAAATTAACACCCACGAAAAATAAAAACCGCACCCTCAACTTCGCAAAAAAC 1001

Qy        556 CACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAG 615
              ||||||  |  |||   || | || | ||||||||| | ||||| | | |||||    |
Db       1000 CACCGTAAAAACCAAAACGATACAAAAACACGACGTATAACTCGAAATACGCCTAAAAAA 941

Qy        616 GATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAG 675
               ||  ||||   | ||    |||||||||   |||||||| |||||||||||||| |
Db        940 AATAAACGAAAAAACGAAAAACCGCTAACGAAACTCCCTCTACGCGCCCCGTCCGCAAAA 881

Qy        676 GCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCA 735
              ||||||||||   |||||  ||  ||||||  |||||  || || |||||| ||| |||
Db        880 ACGCACGTCGAAAATCCCGAACGAACTCCGTAAACGTTAACGATAACGCCGAACGAATCA 821

Qy        736 CGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATC 795
              || ||||| || | ||||||| |||||||  |||||  |||  ||     ||| |||||||
Db        820 CGAACCATAAAAAACGTTCGTACCGCGCGACCCAAAACCGAAATAAAAATTAACCACATC 761

Qy        796 CTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCG 855
              || ||||||| |    | ||||||  ||||  || |||  |||| ||| | |||||||
Db        760 CTACCGCGCTAAAAAAAAAACTAACGAACGCGAACGACCGAACCCAACCGAAACCCACCG 701

Qy        856 CGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCT 915
              ||||  |||   |  | | |||  |||| ||  |||  |||||||| |||||| ||| ||
Db        700 CGATAACGAAAAAAAAATACAAAACGCTACTAAACGAACTCAACAAAACGACTACGTACT 641

Qy        916 ACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGC 975
              ||||||||||  | || ||||||| |  |||| || ||||||| ||||| || ||  | |
Db        640 ACCACCACCTAATACTAACCGTCGATAACTCGACGAACTCGCAAAACCTACGACAAAAAC 581

Qy        976 TGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTG 1035
              | ||||| |||||||| ||  ||||  | ||  || | ||||||| |||||| || |||
Db        580 TACAAAAAACGCGCCAAAAAACGCAAAAACTAACGATATCCACCTACGCCCGACTAACTA 521

Qy       1036 CTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGG 1095
              || | || ||||||||   |||  |||||||||| |||||||| ||||| || ||||
Db        520 CTATACTACGCGACCGAAACCTAACCGCCGACGAACGCGCCGAATTCGAACGACTCTAAA 461

Qy       1096 TGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGG 1155
              |  ||||||||  || |||  |||| ||  || || |||| ||||||  ||||  | ||
Db        460 TAACCTTCTCGAACTACCTAAACCTACTAAAAACGAACATACGACGCGCGCTAAAACTAA 401

Qy       1156 GCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCG 1215
              ||||||||||||||| |||||||||| |||||||  | ||||||  | |  ||  ||
Db        400 ACGCCGCGTTCCCGCTACACGCGCCGCGACGACCGCTAATACGCACAAATATAACTAACG 341

Qy       1216 CCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGG 1275
              |||||||| |||  || |||||||||| | |||||||| ||| || |||||  || |
Db        340 CCTCCTCCGACGTAACGACGCGCGCGCTAAACACCCGCAACCTACGACTCGAAACGAAAA 281

Qy       1276 GCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGA 1335
              |||||||||||||||| |||| ||  | ||  | |||||  |||||||  |   ||| |
Db        280 ACGACTTCGACGTCGCGAACCTACGAAAACTAAAACGCGAAATCCTTCAAATAAACGAAA 221

Qy       1336 TGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGG 1395
              | ||||||||||   | || ||  |||||||  |||||||  |||||  |||  |||| ||
Db        220 TAATCGACAACATAAAAATAAAAATCAACGTACCCCGCTAAACCGTACAAACCCGACAAA 161

Qy       1396 CGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGC 1455
              || ||  |||||| |||||  |||||| ||| |||||  |||||||||| |||| ||||| |
Db        160 CGACGAACGCCGAACTCCTATCCACGATCAACGCCGACCCCTCCTCGATCGTATCCTTAC 101

Qy       1456 AGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCG 1515
              |  | ||||       || ||||||||  ||  ||||  |||||||||||||| |||||
Db        100 AAAAACGCGAAAAAAATTACGACCCCAAAAAAACCCTAACCGCCATCCTTTTCGACGCCG 41

Qy       1516 TGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              | || ||  || || |  ||||| |||| | |||  ||||
Db          40 TACTACTAACGACTATAACCCTAACCGTATACGTAACGAA 1

RESULT 10
US-10-363-483A-27370
; Sequence 27370, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27370
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 27370
US-10-363-483A-27370

  Query Match            19.0%;  Score 594.4;  DB 9;  Length 1061;
  Best Local Similarity   72.5%;  Pred. No. 9.8e-141;
  Matches  769;  Conservative    0;  Mismatches  291;  Indels    0;  Gaps    0;

Qy        496 CTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGC 555
              ||| ||||| | ||  |||||||| |  ||    |||||||||||| |||||||   | |
Db          2 CTTACCCCGAAATTAACACCCACGAAAAATAAAAACCGCACCCTCAACTTCGCAAAAAAC 61

Qy        556 CACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAG 615
              ||||||  |  |||   || | || | ||||||||| | ||||| | | |||||    |
Db         62 CACCGTAAAAACCAAAACGATACAAAAACACGACGTATAACTCGAAATACGCCTAAAAAA 121

Qy        616 GATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAG 675
               ||  ||||   | ||    |||||||||   |||||||| |||||||||||||| |
Db        122 AATAAACGAAAAAACGAAAAACCGCTAACGAAACTCCCTCTACGCGCCCCGTCCGCAAAA 181

Qy        676 GCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCA 735
              ||||||||||   |||||  ||  ||||||  |||||  || || |||||| ||| |||
Db        182 ACGCACGTCGAAAATCCCGAACGAACTCCGTAAACGTTAACGATAACGCCGAACGAATCA 241

Qy        736 CGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATC 795
```

```
Db        242 CGAACCATAAAAAACGTTCGTACCGCGCGACCCAAAACCGAAATAAAAATTAACCACATC 301

Qy        796 CTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCG 855

Db        302 CTACCGCGCTAAAAAAAAAACTAACGAACGCGAACGACCGAACCCAACCGAAACCCACCG 361

Qy        856 CGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCT 915

Db        362 CGATAACGAAAAAAAAATACAAAACGCTACTAAACGAACTCAACAAAACGACTACGTACT 421

Qy        916 ACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGC 975

Db        422 ACCACCACCTAATACTAACCGTCGATAACTCGACGAACTCGCAAAACCTACGACAAAAAC 481

Qy        976 TGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTG 1035

Db        482 TACAAAAAACGCGCCAAAAAACGCAAAAACTAACGATATCCACCTACGCCCGACTAACTA 541

Qy       1036 CTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGG 1095

Db        542 CTATACTACGCGACCGAAACCTAACCGCCGACGAACGCGCCGAATTCGAACGACTCTAAA 601

Qy       1096 TGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGG 1155

Db        602 TAACCTTCTCGAACTACCTAAACCTACTAAAAACGAACATACGACGCGCGCTAAAACTAA 661

Qy       1156 GCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCG 1215

Db        662 ACGCCGCGTTCCCGCTACACGCGCCGCGACGACCGCTAATACGCACAAATATAACTAACG 721

Qy       1216 CCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGG 1275

Db        722 CCTCCTCCGACGTAACGACGCGCGCGCTAAACACCCGCAACCTACGACTCGAAACGAAAA 781

Qy       1276 GCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGA 1335

Db        782 ACGACTTCGACGTCGCGAACCTACGAAAACTAAAACGCGAAATCCTTCAAATAAACGAAA 841

Qy       1336 TGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGG 1395

Db        842 TAATCGACAACATAAAAATAAAAATCAACGTACCCCGCTAAACCGTACAAACCCGACAAA 901

Qy       1396 CGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGC 1455

Db        902 CGACGAACGCCGAACTCCTATCCACGATCAACGCCGACCCCTCCTCGATCGTATCCTTAC 961

Qy       1456 AGGAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCG 1515

Db        962 AAAAACGCGAAAAAAATTACGACCCCAAAAAAACCCTAACCGCCATCCTTTTCGACGCCG 1021

Qy       1516 TGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555

Db       1022 TACTACTAACGACTATAACCCTAACCGTATACGTAACGAA 1061
```

RESULT 11
US-09-925-065A-740956/c
; Sequence 740956, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 740956
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-740956

Query Match 17.9%; Score 559.6; DB 4; Length 560;
Best Local Similarity 99.8%; Pred. No. 5.9e-132;
Matches 559; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy        199 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 258

Db        560 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 501

Qy        259 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 318

Db        500 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 441

Qy        319 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 378

Db        440 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 381

Qy        379 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAA 438

Db        380 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAA 321

Qy        439 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 498

Db        320 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 261

Qy        499 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 558

Db        260 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 201

Qy        559 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 618

Db        200 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 141

Qy        619 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 678

Db        140 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 81

Qy        679 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 738

Db         80 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 21

Qy        739 ACCATGAAGAGCGTTCGTGC 758

Db         20 ACCATGAAGAGCGTTCGTGC 1
```

RESULT 12
US-09-925-065A-769508/c
; Sequence 769508, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769508
;   LENGTH: 557
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-769508

  Query Match             17.8%;  Score 556.6;  DB 4;  Length 557;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-131;
  Matches  556;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        202 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 498

Qy        262 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 438

Qy        322 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 378

Qy        382 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACC 441
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        377 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACC 318

Qy        442 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 258

Qy        502 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 198

Qy        562 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 138

Qy        622 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 78

Qy        682 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 18

Qy        742 ATGAAGAGCGTTCGTGC 758
              |||||||||||||||||
Db         17 ATGAAGAGCGTTCGTGC 1


RESULT 13
US-09-925-065A-736351/c
; Sequence 736351, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 736351
;   LENGTH: 556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-736351

  Query Match             17.8%;  Score 555.6;  DB 4;  Length 556;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-131;
  Matches  555;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 497

Qy        263 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 437

Qy        323 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 377

Qy        383 TGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCC 442
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        376 TGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCC 317

Qy        443 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 257

Qy        503 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 197

Qy        563 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 137

Qy        623 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 77

Qy        683 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 17

Qy        743 TGAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db         16 TGAAGAGCGTTCGTGC 1


RESULT 14
US-09-925-065A-737120
; Sequence 737120, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
```

```
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737120
;   LENGTH: 555
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-737120

  Query Match            17.8%;  Score 554.6;  DB 4;  Length 555;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-130;
  Matches  554;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        204 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 60

Qy        264 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 120

Qy        324 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 180

Qy        384 GATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 443
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        181 GATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 240

Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 300

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 360

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 420

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 480

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 540

Qy        744 GAAGAGCGTTCGTGC 758
              |||||||||||||||
Db        541 GAAGAGCGTTCGTGC 555


RESULT 15
US-10-029-386-6988/c
; Sequence 6988, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6988
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010615.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: AW302149.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q06805, EVALUE 1.90e-01
;   OTHER INFORMATION: NT HIT: gi4507086, EVALUE 7.80e+00
US-10-029-386-6988

  Query Match            16.8%;  Score 524;  DB 6;  Length 524;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-123;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        484 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 465

Qy        544 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 405

Qy        604 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 345

Qy        664 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 285

Qy        724 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 225

Qy        784 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGC 165

Qy        844 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 105

Qy        904 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 45

Qy        964 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1007
              |||||||||||||||||||||||||||||||||||||||||||
Db         44 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1


Search completed: May 11, 2006, 03:56:35
Job time : 2410 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 02:14:07 ; Search time 1415 Seconds
(without alignments)
8998.684 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9312410 seqs, 2039259788 residues

Total number of hits satisfying chosen parameters: 18624820

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq1:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
8: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
11: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
12: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
13: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq4:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
16: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
17: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
18: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq5:*
19: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	559.6	17.9	560	7	US-09-925-065A-740956	Sequence 740956,
c	2	556.6	17.8	557	7	US-09-925-065A-769508	Sequence 769508,
c	3	555.6	17.8	556	7	US-09-925-065A-736351	Sequence 736351,
	4	554.6	17.8	555	7	US-09-925-065A-737120	Sequence 737120,
c	5	180.2	5.8	2306	7	US-09-925-065A-706227	Sequence 706227,
	6	178.6	5.7	414	7	US-09-925-065A-132933	Sequence 132933,
c	7	176.2	5.6	154548	17	US-11-121-086-33	Sequence 33, Appl
c	8	176	5.6	433	11	US-10-301-480-91009	Sequence 91009, A
c	9	176	5.6	433	11	US-10-301-480-91010	Sequence 91010, A
c	10	176	5.6	433	12	US-10-301-480-704418	Sequence 704418,
c	11	176	5.6	433	12	US-10-301-480-704419	Sequence 704419,
	12	175.2	5.6	534	7	US-09-925-065A-64658	Sequence 64658, A
	13	175.2	5.6	534	11	US-10-301-480-165896	Sequence 165896,
	14	175.2	5.6	534	12	US-10-301-480-779305	Sequence 779305,
	15	174.8	5.6	1541	7	US-09-925-065A-89171	Sequence 89171, A
	16	174.8	5.6	1541	7	US-09-925-065A-89172	Sequence 89172, A
	17	174.8	5.6	1541	11	US-10-301-480-190412	Sequence 190412,
	18	174.8	5.6	1541	11	US-10-301-480-190413	Sequence 190413,
	19	174.8	5.6	1541	12	US-10-301-480-803821	Sequence 803821,
	20	174.8	5.6	1541	12	US-10-301-480-803822	Sequence 803822,
c	21	174.8	5.6	121736	18	US-11-114-798-49	Sequence 49, Appl
c	22	174.8	5.6	318488	18	US-11-114-798-58	Sequence 58, Appl
	23	174.4	5.6	1600	7	US-09-925-065A-84740	Sequence 84740, A
	24	174.4	5.6	1600	11	US-10-301-480-185980	Sequence 185980,
	25	174.4	5.6	1600	12	US-10-301-480-799389	Sequence 799389,
	26	174.2	5.6	1691140	17	US-11-091-018-1	Sequence 1, Appli
	27	174	5.6	516	7	US-09-925-065A-489267	Sequence 489267,
	28	174	5.6	550	7	US-09-925-065A-546135	Sequence 546135,
	29	174	5.6	550	12	US-10-301-480-524085	Sequence 524085,
	30	174	5.6	550	12	US-10-301-480-1137494	Sequence 1137494,
	31	174	5.6	150481	17	US-11-112-908-37	Sequence 37, Appl
c	32	174	5.6	155515	17	US-11-112-908-42	Sequence 42, Appl
c	33	174	5.6	159660	17	US-11-112-908-43	Sequence 43, Appl
	34	174	5.6	171162	17	US-11-112-908-38	Sequence 38, Appl
c	35	174	5.6	177623	17	US-11-112-908-41	Sequence 41, Appl
	36	173.6	5.6	386	12	US-10-301-480-229094	Sequence 229094,
	37	173.6	5.6	386	12	US-10-301-480-842503	Sequence 842503,
c	38	173.6	5.6	717	7	US-09-925-065A-939937	Sequence 939937,
c	39	173.6	5.6	717	7	US-09-925-065A-955036	Sequence 955036,
c	40	173.6	5.6	908	12	US-10-301-480-609462	Sequence 609462,
c	41	173.6	5.6	908	12	US-10-301-480-1222871	Sequence 1222871,
	42	173.4	5.6	1549	7	US-09-925-065A-726968	Sequence 726968,
c	43	173.4	5.6	101046	10	US-10-995-561-13330	Sequence 13330, A
	44	173.2	5.5	679	12	US-10-301-480-531347	Sequence 531347,
	45	173.2	5.5	679	12	US-10-301-480-1144756	Sequence 1144756,

ALIGNMENTS

RESULT 1
US-09-925-065A-740956/c
; Sequence 740956, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 740956
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-740956

Query Match 17.9%; Score 559.6; DB 7; Length 560;
Best Local Similarity 99.8%; Pred. No. 0.017;
Matches 559; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 258
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 560 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 501

```
Qy        259 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 441

Qy        319 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 381

Qy        379 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAA 438
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        380 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAA 321

Qy        439 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 261

Qy        499 GCCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 201

Qy        559 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 141

Qy        619 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 81

Qy        679 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 21

Qy        739 ACCATGAAGAGCGTTCGTGC 758
              ||||||||||||||||||||
Db         20 ACCATGAAGAGCGTTCGTGC 1
```

```
RESULT 2
US-09-925-065A-769508/c
; Sequence 769508, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769508
;   LENGTH: 557
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-769508

  Query Match            17.8%;  Score 556.6;  DB 7;  Length 557;
  Best Local Similarity   99.8%;  Pred. No. 0.02;
  Matches  556;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        202 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 498
```

```
Qy        262 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 438

Qy        322 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 378

Qy        382 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACC 441
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        377 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACC 318

Qy        442 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 258

Qy        502 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 198

Qy        562 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 138

Qy        622 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 78

Qy        682 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 18

Qy        742 ATGAAGAGCGTTCGTGC 758
              |||||||||||||||||
Db         17 ATGAAGAGCGTTCGTGC 1
```

```
RESULT 3
US-09-925-065A-736351/c
; Sequence 736351, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 736351
;   LENGTH: 556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-736351

  Query Match            17.8%;  Score 555.6;  DB 7;  Length 556;
  Best Local Similarity   99.8%;  Pred. No. 0.02;
  Matches  555;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        203 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        556 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 497

Qy        263 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 437

Qy        323 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 377

Qy        383 TGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCC 442
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        376 TGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCC 317

Qy        443 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 257

Qy        503 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 197

Qy        563 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 137

Qy        623 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 77

Qy        683 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 17

Qy        743 TGAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db         16 TGAAGAGCGTTCGTGC 1
```

RESULT 4
US-09-925-065A-737120
; Sequence 737120, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737120
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-737120

Query Match 17.8%; Score 554.6; DB 7; Length 555;
Best Local Similarity 99.8%; Pred. No. 0.021;
Matches 554; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy        204 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 60

Qy        264 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 120

Qy        324 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 180

Qy        384 GATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 443
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        181 GATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 240

Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 300

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 360

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 420

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 480

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 540

Qy        744 GAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db        541 GAAGAGCGTTCGTGC 555
```

RESULT 5
US-09-925-065A-706227/c
; Sequence 706227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 706227
; LENGTH: 2306
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-706227

Query Match 5.8%; Score 180.2; DB 7; Length 2306;
Best Local Similarity 84.6%; Pred. No. 5.5e+03;
Matches 214; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

```
Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
Db        318 CCCCGCACTTTGGGAGGCCGAGGCAGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGC 259

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
Db        258 CTGGCCAACATGGTGAAACCCCGTCTCTACT-AAAAATACAAAAATTAGCCAGGCATGGT 200

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
Db        199 CGTGCATGCCTCTAATCCCAGCTACTCAGGAGGCTGAGGTGGGAGAATCGCTTGAATCCA 140

Qy       3050 GGAGGTGGAGGTTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGC 3109
Db        139 GGAGGTGGAGGTTGCAGTGAGCAAGACTGCACCACTGCACTCCAGGATGGGCAACACAGC 80

Qy       3110 AAGACTCTGTCTC 3122
Db          79 AAGACTCTGTCTC 67
```

RESULT 6
US-09-925-065A-132933
; Sequence 132933, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132933
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-132933

Query Match 5.7%; Score 178.6; DB 7; Length 414;
Best Local Similarity 84.2%; Pred. No. 2.2e+04;
Matches 213; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

```
Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
Db        139 CCCAGCACTTTGGGAGACTGAGGTGGGTGGATCACTTAAGGCCAGGAGTTTGAGACCAGC 198

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
Db        199 CTGGTCAACAGAGCAAAACCCCATCTCTACTAAAAAATACAAAAATTAGCTGGGCATGGT 258

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
Db        259 GGCGCAAGCCTATAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCG 318

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
Db        319 GGAGGCGGAGGTTGCAGTGAGCCAATATCGTACCACTGCACTCCAGCCTGGGCAACAGAG 378

Qy       3109 CAAGACTCTGTCT 3121
Db        379 CAAAACTCTGTCT 391
```

RESULT 7
US-11-121-086-33/c
; Sequence 33, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
; LENGTH: 154548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-33

Query Match 5.6%; Score 176.2; DB 17; Length 154548;
Best Local Similarity 86.8%; Pred. No. 2.5e+02;
Matches 204; Conservative 1; Mismatches 29; Indels 1; Gaps 1;

```
Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAAC 2948
Db      25277 GAGGCAGGCGGATCAACTGAGGTCAGGAGTTCAAGACCAGCCTGACCAACATGGTGAAAC 25218

Qy       2949 CCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCC 3008
Db      25217 CCCGTCTCTACTAAAAAATATGAAAATTAGCTAGGCRTGGTGGTGCACGCCTGTAGTCCC 25158

Qy       3009 AGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATG 3068
Db      25157 AGCTACTCTGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGTTGCAGTG 25098

Qy       3069 AG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      25097 AGCCAAGATTGCGCCATTGCACTCCAGCCTGGGCGACAGAGTGAGACTCTGTCTC 25043
```

RESULT 8
US-10-301-480-91009/c
; Sequence 91009, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91009
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-91009

Query Match 5.6%; Score 176; DB 11; Length 433;
Best Local Similarity 83.1%; Pred. No. 2.3e+04;
Matches 211; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

```
Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
```

```
Db        307 CCCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAGGTCAGGAGTTTGAGACCAGC 248

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||||||||| ||||||| |||||| |||||||    ||||||| || |||||||||
Db        247 CTGGCCAACATAGTGAAACCCCATCTCTTCTAAAAATACAAAAAATTAGCTGGGCATGGT 188

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             |||:|| ||||||  |||| ||||||| ||||| |||||||||||| || ||||||  ||
Db        187 GGCSCACGCCTGTAATCCCTGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCA 128

Qy       3050 GGAGGTGGAGGTTGCAATG-AGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             ||||||||||||||| || ||| ||||  |||||||||||| ||||||||||||||||||
Db        127 GGAGGTGGAGGTTGCAGTGAAGCGAGATCATGCCACTGCACTGCAGCCTGGGCAACAGAG 68

Qy       3109 CAAGACTCTGTCTC 3122
             ||||||||  ||||
Db         67 CAAGACTCCATCTC 54


RESULT 9
US-10-301-480-91010/c
; Sequence 91010, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91010
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-91010

  Query Match             5.6%;  Score 176;  DB 11;  Length 433;
  Best Local Similarity   83.1%;  Pred. No. 2.3e+04;
  Matches  211;  Conservative    1;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             |||  |  |||       |||||| || |||||||||||| |||||||| |||||||:|
Db        307 CCCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAGGTCAGGAGTTTGAGACCARC 248

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||||||||| ||||||| |||||| |||||||    ||||||| || |||||||||
Db        247 CTGGCCAACATAGTGAAACCCCATCTCTTCTAAAAATACAAAAAATTAGCTGGGCATGGT 188

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             |||||| ||||||  |||| ||||||| ||||| |||||||||||| || ||||||  ||
Db        187 GGCGCACGCCTGTAATCCCTGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCA 128

Qy       3050 GGAGGTGGAGGTTGCAATG-AGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             ||||||||||||||| || ||| ||||  |||||||||||| ||||||||||||||||||
Db        127 GGAGGTGGAGGTTGCAGTGAAGCGAGATCATGCCACTGCACTGCAGCCTGGGCAACAGAG 68

Qy       3109 CAAGACTCTGTCTC 3122
             ||||||||  ||||
Db         67 CAAGACTCCATCTC 54


RESULT 10
US-10-301-480-704418/c
; Sequence 704418, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 704418
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-704418

  Query Match             5.6%;  Score 176;  DB 12;  Length 433;
  Best Local Similarity   83.1%;  Pred. No. 2.3e+04;
  Matches  211;  Conservative    1;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             |||  |  |||       |||||| || |||||||||||| |||||||| |||||||:|
Db        307 CCCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAGGTCAGGAGTTTGAGACCARC 248

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             ||||||||||||| ||||||| |||||| |||||||    ||||||| || |||||||||
Db        247 CTGGCCAACATAGTGAAACCCCATCTCTTCTAAAAATACAAAAAATTAGCTGGGCATGGT 188

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             |||:|| ||||||  |||| ||||||| ||||| |||||||||||| || ||||||  ||
Db        187 GGCSCACGCCTGTAATCCCTGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCA 128

Qy       3050 GGAGGTGGAGGTTGCAATG-AGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             ||||||||||||||| || ||| ||||  |||||||||||| ||||||||||||||||||
Db        127 GGAGGTGGAGGTTGCAGTGAAGCGAGATCATGCCACTGCACTGCAGCCTGGGCAACAGAG 68

Qy       3109 CAAGACTCTGTCTC 3122
             ||||||||  ||||
Db         67 CAAGACTCCATCTC 54


RESULT 11
US-10-301-480-704419/c
; Sequence 704419, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 704419
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-704419

  Query Match             5.6%;  Score 176;  DB 12;  Length 433;
  Best Local Similarity   83.1%;  Pred. No. 2.3e+04;
  Matches  211;  Conservative    1;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
             |||  |  |||       |||||| || |||||||||||| |||||||| |||||||:|
Db        307 CCCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAGGTCAGGAGTTTGAGACCARC 248
```

```
Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
             |||||||||||| ||||||| |||||| |||||||     ||||||| || |||||||||
Db        247 CTGGCCAACATAGTGAAACCCCATCTCTTCTAAAAATACAAAAAATTAGCTGGGCATGGT 188

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
             |||||| ||||||  |||| |||||||| ||||| |||||||||||| || ||||||  ||
Db        187 GGCGCACGCCTGTAATCCCTGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCA 128

Qy       3050 GGAGGTGGAGGTTGCAATG-AGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
             ||||||||||||||||  || ||| ||||  |||||||||||| ||||||||||||||||
Db        127 GGAGGTGGAGGTTGCAGTGAAGCGAGATCATGCCACTGCACTGCAGCCTGGGCAACAGAG 68

Qy       3109 CAAGACTCTGTCTC 3122
             ||||||||  ||||
Db          67 CAAGACTCCATCTC 54
```

```
RESULT 12
US-09-925-065A-64658
; Sequence 64658, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64658
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-64658

  Query Match             5.6%;  Score 175.2;  DB 7;  Length 534;
  Best Local Similarity   86.8%;  Pred. No. 2e+04;
  Matches  203;  Conservative    1;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACC 2949
             |||||||||||||||||||||| ||||| ||| |||||||||||||||||| | ||||||
Db        130 AGGCAGGTGGATCACCTGAGGTCAGGAATTCAAGACCAGCCTGGCCAACATGGTGAAACC 189

Qy       2950 CGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCA 3009
             |  |||||||||| ||||| ||||||| |||  |||||||||||||||||||  |||||
Db        190 CTGTCTCTACTAACAAATACAAAAATTAGCCATGCATGGTGGCGCATGCCTGTAATCCCA 249

Qy       3010 GCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATGA 3069
             |||||: |||||| ||||||||||||   |||||||    |||||||||||||||| |||
Db        250 GCTACYAGGGAGGCTGAGGCAGGAGAACTGCTTGAACCTGGGAGGTGGAGGTTGCAGTGA 309

Qy       3070 G-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             | ||||||  |||||||||| ||||| |||||| ||||||||||||||||||||
Db        310 GCCAAGATCATGCCACTGCAATCCAGGCTGGGCGACAGAGCAAGACTCTGTCTC 363
```

```
RESULT 13
US-10-301-480-165896
; Sequence 165896, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165896
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-165896

  Query Match             5.6%;  Score 175.2;  DB 11;  Length 534;
  Best Local Similarity   86.8%;  Pred. No. 2e+04;
  Matches  203;  Conservative    1;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACC 2949
             |||||||||||||||||||||| ||||| ||| |||||||||||||||||| | ||||||
Db        130 AGGCAGGTGGATCACCTGAGGTCAGGAATTCAAGACCAGCCTGGCCAACATGGTGAAACC 189

Qy       2950 CGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCA 3009
             |  |||||||||| ||||| ||||||| |||  |||||||||||||||||||  |||||
Db        190 CTGTCTCTACTAACAAATACAAAAATTAGCCATGCATGGTGGCGCATGCCTGTAATCCCA 249

Qy       3010 GCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATGA 3069
             |||||: |||||| ||||||||||||   |||||||    |||||||||||||||| |||
Db        250 GCTACYAGGGAGGCTGAGGCAGGAGAACTGCTTGAACCTGGGAGGTGGAGGTTGCAGTGA 309

Qy       3070 G-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             | ||||||  |||||||||| ||||| |||||| ||||||||||||||||||||
Db        310 GCCAAGATCATGCCACTGCAATCCAGGCTGGGCGACAGAGCAAGACTCTGTCTC 363
```

```
RESULT 14
US-10-301-480-779305
; Sequence 779305, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 779305
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-779305

  Query Match             5.6%;  Score 175.2;  DB 12;  Length 534;
  Best Local Similarity   86.8%;  Pred. No. 2e+04;
  Matches  203;  Conservative    1;  Mismatches   29;  Indels    1;  Gaps    1;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACC 2949
             |||||||||||||||||||||| ||||| ||| |||||||||||||||||| | ||||||
Db        130 AGGCAGGTGGATCACCTGAGGTCAGGAATTCAAGACCAGCCTGGCCAACATGGTGAAACC 189

Qy       2950 CGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCA 3009
             |  |||||||||| ||||| ||||||| |||  |||||||||||||||||||  |||||
Db        190 CTGTCTCTACTAACAAATACAAAAATTAGCCATGCATGGTGGCGCATGCCTGTAATCCCA 249
```

```
Qy       3010 GCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGGTTGCAATGA 3069
              |||||: |||||| |||||||||||||   |||||||     |||||||||||||||| |||
Db        250 GCTACYAGGGAGGCTGAGGCAGGAGAACTGCTTGAACCTGGGAGGTGGAGGTTGCAGTGA 309

Qy       3070 G-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              | ||||||  |||||||||| ||||| |||||| ||||||||||||||||||||
Db        310 GCCAAGATCATGCCACTGCAATCCAGGCTGGGCGACAGAGCAAGACTCTGTCTC 363


RESULT 15
US-09-925-065A-89171
; Sequence 89171, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89171
;   LENGTH: 1541
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-89171

  Query Match             5.6%;  Score 174.8;  DB 7;  Length 1541;
  Best Local Similarity   83.1%;  Pred. No. 9.1e+03;
  Matches  211;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;

Qy       2870 CCCCACTTTTTTCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGC 2929
              |||  |  |||        |||||||||||||||||||||| ||||||||||||||||
Db        533 CCCAGCAGTTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAGT 592

Qy       2930 CTGGCCAACATAGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGT 2989
              ||||||| || || ||||||| ||||| |||||||     ||||||| || |||  ||||
Db        593 CTGGCCAGCACAGTGAAACCCCATCTCCACTAAAACTACAAAAAATTAGCTGGGTGTGGT 652

Qy       2990 GGCGCATGCCTGTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCA 3049
              ||||  |||||||  |||||||||||||||||| ||||||| |||| ||| |||||  ||
Db        653 GGCGGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCACGAGAATCGTTTGAACCCA 712

Qy       3050 GGAGGTGGAGGTTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
              |||||||||||||||| |||| ||| ||||||||||||||||||||||||||||||| ||
Db        713 GGAGGTGGAGGTTGCACTGAGCCAAAATTGTGCCACTGCACTCCAGCCTGGGCAACATAG 772

Qy       3109 CAAGACTCTGTCTC 3122
              ||||| ||| ||||
Db        773 CAAGATTCTATCTC 786


Search completed: May 11, 2006, 05:51:18
Job time : 1416 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 03:56:47 ; Search time 15378 Seconds
(without alignments)
11540.204 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 1

Total number of hits satisfying chosen parameters: 11765110

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3122	100.0	3122	6	CQ841455	CQ841455 Sequence
	2	3122	100.0	3122	8	AK124499	AK124499 Homo sapi
c	3	2840	91.0	191925	8	AC008736	AC008736 Homo sapi
c	4	2840	91.0	216441	14	AC027340	AC027340 Homo sapi
	5	2445	78.3	186115	8	AC008474	AC008474 Homo sapi
	6	1952	62.5	2327	6	CQ850499	CQ850499 Sequence
	7	1952	62.5	2327	8	AK127646	AK127646 Homo sapi
	8	657	21.0	708	6	CQ748891	CQ748891 Sequence
	9	646	20.7	655	8	HUMZD85E06	AF086461 Homo sapi
	10	606	19.4	708	6	CQ746021	CQ746021 Sequence
	11	392	12.6	216441	14	AC027340	AC027340 Homo sapi
c	12	144	4.6	449	10	G37408	G37408 SHGC-57769
	13	60	1.9	60	6	CQ553695	CQ553695 Sequence
	14	53	1.7	24495	8	CR762434	CR762434 Human DNA
c	15	53	1.7	42398	8	D84401	D84401 Homo sapien
	16	53	1.7	49878	8	CR547129	CR547129 Human DNA
	17	53	1.7	62449	14	AC061985	AC061985 Homo sapi
c	18	53	1.7	69808	8	AC026771	AC026771 Homo sapi
	19	53	1.7	70621	8	AL713971	AL713971 Human DNA
	20	53	1.7	90476	8	AC015910	AC015910 Homo sapi
	21	53	1.7	100000	8	AP000208	AP000208 Homo sapi
	22	53	1.7	109523	8	AL844527	AL844527 Human DNA
c	23	53	1.7	110130	8	AL360268	AL360268 Human DNA
	24	53	1.7	113919	8	AP000247	AP000247 Homo sapi
	25	53	1.7	120825	8	HSJ733M16	AL109627 Human DNA
c	26	53	1.7	134184	8	AC095044	AC095044 Homo sapi
	27	53	1.7	157686	8	BS000110	BS000110 Pan trogl
	28	53	1.7	158033	8	AL645940	AL645940 Human DNA
	29	53	1.7	171051	8	AC022960	AC022960 Homo sapi
c	30	53	1.7	175737	8	HS1033B10	AL031228 Human DNA
	31	53	1.7	179462	14	AC021102	AC021102 Homo sapi
	32	53	1.7	185000	14	AC007332	AC007332 Homo sapi
	33	53	1.7	189181	14	AP001333	AP001333 Homo sapi
	34	53	1.7	190092	14	AC120502	AC120502 Lemur cat
	35	53	1.7	203583	8	AC092437	AC092437 Homo sapi
	36	53	1.7	204373	14	CR847805	CR847805 Homo sapi
	37	53	1.7	208220	8	AC009107	AC009107 Homo sapi
	38	53	1.7	218872	14	AC092434	AC092434 Homo sapi
	39	53	1.7	229090	14	CR847784	CR847784 Homo sapi
	40	53	1.7	323185	14	CR753824	CR753824 Homo sapi
	41	53	1.7	340000	8	AP001710	AP001710 Homo sapi
	42	53	1.7	349980	6	CS039425	CS039425 Sequence
c	43	52	1.7	381	6	BD027540	BD027540 Sequence
c	44	52	1.7	381	6	AX887930	AX887930 Sequence
c	45	52	1.7	428	10	AB135735	AB135735 Homo sapi
c	46	52	1.7	690	10	BV640976	BV640976 S217P6013
	47	52	1.7	762	10	BV533342	BV533342 G591P6262
	48	52	1.7	770	10	BV536183	BV536183 G591P6314
	49	52	1.7	773	10	BV588891	BV588891 G591P6387
	50	52	1.7	775	10	BV492390	BV492390 S221P6065
c	51	52	1.7	810	10	BV523612	BV523612 G591P6007
	52	52	1.7	915	10	BV465101	BV465101 G591P6193
c	53	52	1.7	22738	8	AB014078	AB014078 Homo sapi
	54	52	1.7	23481	8	BX248096	BX248096 Human DNA
c	55	52	1.7	25057	8	AC087649	AC087649 Homo sapi
	56	52	1.7	25242	8	AL356788	AL356788 Human DNA
	57	52	1.7	25972	8	CR936916	CR936916 Human DNA
	58	52	1.7	28276	8	CR382280	CR382280 Human DNA
c	59	52	1.7	32906	8	CR391992	CR391992 Human DNA
c	60	52	1.7	38173	8	AC093235	AC093235 Homo sapi
	61	52	1.7	39372	8	HSN121E8	Z78421 Human DNA s
c	62	52	1.7	39372	8	HSN121E8	Z78421 Human DNA s
c	63	52	1.7	39596	14	AL445257	AL445257 Homo sapi
	64	52	1.7	41372	8	AL359082	AL359082 Human DNA
	65	52	1.7	42572	8	AC004602	AC004602 Homo sapi
	66	52	1.7	44449	8	AC016627	AC016627 Homo sapi
	67	52	1.7	44783	8	CR936927	CR936927 Human DNA
c	68	52	1.7	46201	8	AC006139	AC006139 Homo sapi
c	69	52	1.7	46509	8	AC108747	AC108747 Homo sapi
	70	52	1.7	47462	8	CR759912	CR759912 Human DNA
	71	52	1.7	51090	14	AL390837	AL390837 Homo sapi
	72	52	1.7	59744	14	AC103989	AC103989 Homo sapi
c	73	52	1.7	62761	8	AL499606	AL499606 Human DNA
	74	52	1.7	64056	8	AC092080	AC092080 Homo sapi
c	75	52	1.7	66887	8	AF411057	AF411057 Homo sapi
c	76	52	1.7	68959	8	HSJ735P11	AL049735 Human DNA
c	77	52	1.7	70393	8	AC073488	AC073488 Homo sapi
	78	52	1.7	70448	8	CR388372	CR388372 Human DNA
	79	52	1.7	73597	14	AC109598	AC109598 Homo sapi
c	80	52	1.7	75525	8	AL391561	AL391561 Human DNA
	81	52	1.7	80117	8	AC055813	AC055813 Homo sapi
	82	52	1.7	82724	8	BX927220	BX927220 Human DNA
	83	52	1.7	92139	6	AR242855	AR242855 Sequence
	84	52	1.7	92139	6	AX384907	AX384907 Sequence
c	85	52	1.7	92442	14	CR751233	CR751233 Homo sapi
c	86	52	1.7	93790	8	AF397423	AF397423 Homo sapi
	87	52	1.7	96217	8	AL645729	AL645729 Human DNA
	88	52	1.7	97630	8	AL844213	AL844213 Human DNA
c	89	52	1.7	101769	14	AC093203	AC093203 Homo sapi
c	90	52	1.7	102311	8	AC005377	AC005377 Homo sapi
c	91	52	1.7	105902	8	AL359699	AL359699 Human DNA

	92	52	1.7	105956	8	AC074136 AC074136 Homo sapi
c	93	52	1.7	106539	8	AF225899 AF225899 Homo sapi
c	94	52	1.7	110000	8	BA000025_14 Continuation (15 o
c	95	52	1.7	110000	8	BA000041_09 Continuation (10 o
c	96	52	1.7	110000	8	BA000041_10 Continuation (11 o
c	97	52	1.7	110028	8	HS1164I10 AL049537 Human DNA
	98	52	1.7	110042	8	AL590788 AL590788 Human DNA
c	99	52	1.7	111051	14	AF322449 AF322449 Homo sapi
c	100	52	1.7	116888	8	AC114481 AC114481 Homo sapi
	101	52	1.7	117899	8	AC016590 AC016590 Homo sapi
c	102	52	1.7	120723	8	AC104051 AC104051 Homo sapi
c	103	52	1.7	128540	14	AC025473 AC025473 Homo sapi
c	104	52	1.7	130146	8	AL590489 AL590489 Human DNA
	105	52	1.7	130572	14	AC024096 AC024096 Homo sapi
	106	52	1.7	130984	14	AC022766 AC022766 Homo sapi
c	107	52	1.7	134465	8	AC100793 AC100793 Homo sapi
	108	52	1.7	135805	14	AC069511 AC069511 Homo sapi
c	109	52	1.7	136385	8	AC006006 AC006006 Homo sapi
c	110	52	1.7	137845	14	AC010367 AC010367 Homo sapi
	111	52	1.7	138573	8	AL954211 AL954211 Pan trogl
	112	52	1.7	141633	8	AC073542 AC073542 Homo sapi
c	113	52	1.7	143969	8	AL357552 AL357552 Human DNA
	114	52	1.7	144542	14	AC027146 AC027146 Homo sapi
	115	52	1.7	146199	14	AC110281 AC110281 Homo sapi
	116	52	1.7	146376	8	AC009247 AC009247 Homo sapi
c	117	52	1.7	146740	8	CNS01DTX AL132989 Human chr
c	118	52	1.7	148049	8	HSA300188 AJ300188 Homo sapi
c	119	52	1.7	149679	8	AC002351 AC002351 Homo sapi
	120	52	1.7	149776	8	AC012070 AC012070 Homo sapi
	121	52	1.7	150159	14	AC018791 AC018791 Homo sapi
	122	52	1.7	150437	6	CS086345 CS086345 Sequence
	123	52	1.7	150846	8	AL138724 AL138724 Human DNA
	124	52	1.7	151228	8	AL662800 AL662800 Human DNA
c	125	52	1.7	151970	14	AC040905 AC040905 Homo sapi
c	126	52	1.7	152040	14	AC139547 AC139547 Homo sapi
	127	52	1.7	152492	8	AC079863 AC079863 Homo sapi
c	128	52	1.7	157135	14	AC064817 AC064817 Homo sapi
	129	52	1.7	160929	8	AC008677 AC008677 Homo sapi
c	130	52	1.7	161970	14	AC027706 AC027706 Homo sapi
	131	52	1.7	162948	8	AL662822 AL662822 Human DNA
c	132	52	1.7	163157	8	AC108670 AC108670 Homo sapi
	133	52	1.7	163280	14	AC141308 AC141308 Homo sapi
	134	52	1.7	163432	14	AC161476 AC161476 Pan trogl
	135	52	1.7	163437	8	AC092606 AC092606 Homo sapi
c	136	52	1.7	163662	14	AC096508 AC096508 Homo sapi
c	137	52	1.7	164429	14	AC150837 AC150837 Papio anu
c	138	52	1.7	165311	8	AC097061 AC097061 Homo sapi
c	139	52	1.7	165972	14	AC135854 AC135854 Homo sapi
	140	52	1.7	166484	14	AC026830 AC026830 Homo sapi
	141	52	1.7	167780	8	AL844892 AL844892 Human DNA
	142	52	1.7	167920	8	AC109635 AC109635 Homo sapi
	143	52	1.7	168734	8	AC007371 AC007371 Homo sapi
	144	52	1.7	168793	8	AC025787 AC025787 Homo sapi
c	145	52	1.7	169059	14	AC009790 AC009790 Homo sapi
c	146	52	1.7	169893	14	AC021165 AC021165 Homo sapi
c	147	52	1.7	170232	8	AL135903 AL135903 Human DNA
	148	52	1.7	171058	8	AC073068 AC073068 Homo sapi
	149	52	1.7	171849	8	AC010616 AC010616 Homo sapi
c	150	52	1.7	172307	8	AL590609 AL590609 Human DNA
	151	52	1.7	172759	8	AC025589 AC025589 Homo sapi
c	152	52	1.7	172915	14	AC010160 AC010160 Homo sapi
	153	52	1.7	174034	8	AC020908 AC020908 Homo sapi
c	154	52	1.7	174521	14	AC127468 AC127468 Papio anu
c	155	52	1.7	176343	8	CNS01DX3 AL139021 Human chr
	156	52	1.7	176871	14	AC150824 AC150824 Callithri
c	157	52	1.7	177744	14	AC073954 AC073954 Homo sapi
	158	52	1.7	177773	8	AC010761 AC010761 Homo sapi
c	159	52	1.7	178168	14	AC012350 AC012350 Homo sapi
c	160	52	1.7	181792	8	AC093126 AC093126 Papio anu
c	161	52	1.7	182152	14	AC067929 AC067929 Homo sapi
	162	52	1.7	182892	8	AC034244 AC034244 Homo sapi
c	163	52	1.7	183556	14	AC019059 AC019059 Homo sapi
c	164	52	1.7	184252	14	AC139565 AC139565 Homo sapi
	165	52	1.7	186415	8	AC008675 AC008675 Homo sapi
c	166	52	1.7	186418	8	AC018757 AC018757 Homo sapi
	167	52	1.7	186448	14	AC034129 AC034129 Homo sapi
	168	52	1.7	188596	8	AC120045 AC120045 Homo sapi
	169	52	1.7	188640	14	AC146153 AC146153 Pan trogl
	170	52	1.7	191037	8	AC091561 AC091561 Homo sapi
c	171	52	1.7	191234	14	AC018681 AC018681 Homo sapi
c	172	52	1.7	191540	8	AC144780 AC144780 Pan trogl
	173	52	1.7	191866	14	AC068676 AC068676 Homo sapi
c	174	52	1.7	191924	14	AC091440 AC091440 Homo sapi
	175	52	1.7	192328	8	AL954212 AL954212 Pan trogl
	176	52	1.7	192505	8	AL353136 AL353136 Human DNA
c	177	52	1.7	193351	14	AC110601 AC110601 Homo sapi
c	178	52	1.7	193952	14	AC141263 AC141263 Homo sapi
	179	52	1.7	194189	8	AC103705 AC103705 Homo sapi
	180	52	1.7	195389	8	AP002026 AP002026 Homo sapi
c	181	52	1.7	196622	8	AC125238 AC125238 Homo sapi
	182	52	1.7	196623	8	AC099058 AC099058 Homo sapi
	183	52	1.7	198105	14	AC026160 AC026160 Homo sapi
c	184	52	1.7	200000	14	AC008108 AC008108 Homo sapi
c	185	52	1.7	200853	8	CNS01DSQ AL121839 Human chr
	186	52	1.7	201167	8	AL162591 AL162591 Human DNA
	187	52	1.7	202539	8	AC090559 AC090559 Homo sapi
	188	52	1.7	202971	14	AC151042 AC151042 Callithri
c	189	52	1.7	204917	8	AC008040 AC008040 Homo sapi
c	190	52	1.7	205283	8	AC073916 AC073916 Homo sapi
	191	52	1.7	205736	14	AC023560 AC023560 Homo sapi
c	192	52	1.7	207661	8	AC015884 AC015884 Homo sapi
c	193	52	1.7	211812	14	AC027205 AC027205 Homo sapi
c	194	52	1.7	213947	14	AC135988 AC135988 Homo sapi
	195	52	1.7	215780	14	AC135989 AC135989 Homo sapi
c	196	52	1.7	234053	8	AC002429 AC002429 Homo sapi
c	197	52	1.7	256000	14	AC145896 AC145896 Pan trogl
	198	52	1.7	279430	14	AC145065 AC145065 Pan trogl
	199	52	1.7	349980	6	CS039416 CS039416 Sequence
c	200	51	1.6	715	10	BV634964 BV634964 S217P6621
c	201	51	1.6	14814	8	AL357035 AL357035 Human DNA
	202	51	1.6	34869	14	CR318623 CR318623 Homo sapi
c	203	51	1.6	68003	14	BX323853 BX323853 Homo sapi
	204	51	1.6	71173	8	AC069148 AC069148 Homo sapi
c	205	51	1.6	86882	8	BX649553 BX649553 Human DNA
c	206	51	1.6	93663	8	AC022409 AC022409 Homo sapi
c	207	51	1.6	105392	8	AC073215 AC073215 Homo sapi
c	208	51	1.6	125403	8	AC008929 AC008929 Homo sapi
	209	51	1.6	126682	8	AC008649 AC008649 Homo sapi
	210	51	1.6	148193	8	AC027347 AC027347 Homo sapi
	211	51	1.6	148782	8	AP006213 AP006213 Homo sapi
	212	51	1.6	153998	14	AC013329 AC013329 Homo sapi
c	213	51	1.6	154625	14	CR936360 CR936360 Homo sapi
	214	51	1.6	154754	8	AC064836 AC064836 Homo sapi
c	215	51	1.6	159392	14	AC148829 AC148829 Pan trogl
c	216	51	1.6	163246	14	AC068401 AC068401 Homo sapi
c	217	51	1.6	165433	8	AC146088 AC146088 Pan trogl
c	218	51	1.6	169627	8	AL136984 AL136984 Human DNA
c	219	51	1.6	174034	8	AC020908 AC020908 Homo sapi
	220	51	1.6	174428	8	AC005821 AC005821 Homo sapi
c	221	51	1.6	175501	8	AC008427 AC008427 Homo sapi
	222	51	1.6	175994	8	AC079460 AC079460 Homo sapi
	223	51	1.6	177777	14	AC068862 AC068862 Homo sapi
c	224	51	1.6	178665	14	AP003483 AP003483 Homo sapi
	225	51	1.6	180133	8	AL138885 AL138885 Human DNA
	226	51	1.6	184133	8	AC103559 AC103559 Homo sapi
	227	51	1.6	184289	8	AL359878 AL359878 Human DNA
c	228	51	1.6	184569	8	AC010583 AC010583 Homo sapi
	229	51	1.6	186239	8	AC123982 AC123982 Pan trogl
	230	51	1.6	190708	8	AC005261 AC005261 Homo sapi
	231	51	1.6	191311	14	AC091486 AC091486 Homo sapi
c	232	51	1.6	192826	8	AC113266 AC113266 Papio anu
	233	51	1.6	194296	8	AL354864 AL354864 Human DNA
c	234	51	1.6	195766	8	CNS01DVC AL135744 Human chr
c	235	51	1.6	199827	14	AC026373 AC026373 Homo sapi
	236	51	1.6	219122	8	AC080038 AC080038 Homo sapi
	237	51	1.6	227137	8	AC098481 AC098481 Homo sapi

c	238	50	1.6	255	6	BD048209 BD048209 Sequence
c	239	50	1.6	255	6	AX912676 AX912676 Sequence
c	240	50	1.6	686	10	BV664432 BV664432 S215P6072
	241	50	1.6	726	10	BV617582 BV617582 S217P6943
c	242	50	1.6	730	10	BV622197 BV622197 S216P6202
c	243	50	1.6	817	10	BV668130 BV668130 S217P6025
	244	50	1.6	911	10	BV632403 BV632403 S215P6040
	245	50	1.6	1001	6	CQ853998 CQ853998 Sequence
	246	50	1.6	1258	8	BC030749 BC030749 Homo sapi
c	247	50	1.6	2188	8	AK025471 AK025471 Homo sapi
c	248	50	1.6	3597	8	HUMAPOLIAB L07900 Human apoli
c	249	50	1.6	3625	8	AK130932 AK130932 Homo sapi
	250	50	1.6	8705	6	AR567482 AR567482 Sequence
	251	50	1.6	8705	6	AR569796 AR569796 Sequence
	252	50	1.6	8705	6	AX277531 AX277531 Sequence
	253	50	1.6	8705	6	AX418095 AX418095 Sequence
c	254	50	1.6	15169	8	AC134915 AC134915 Homo sapi
	255	50	1.6	23536	8	AC002120 AC002120 Homo sapi
	256	50	1.6	35865	8	AC011552 AC011552 Homo sapi
c	257	50	1.6	40338	14	AC147414 AC147414 Homo sapi
c	258	50	1.6	48723	8	AL645992 AL645992 Human DNA
	259	50	1.6	51770	8	AL590490 AL590490 Human DNA
c	260	50	1.6	52368	8	AC010621 AC010621 Homo sapi
	261	50	1.6	55672	14	AC103991 AC103991 Homo sapi
c	262	50	1.6	62164	8	AL670729 AL670729 Human DNA
c	263	50	1.6	63437	8	HSJ964N17 AL050311 Human DNA
	264	50	1.6	64475	8	AF283320S2 AF283321 Homo sapi
c	265	50	1.6	64989	14	AC024899 AC024899 Homo sapi
c	266	50	1.6	65268	14	AC099847 AC099847 Homo sapi
	267	50	1.6	65958	8	AC105413 AC105413 Homo sapi
	268	50	1.6	69926	8	AL593851 AL593851 Human DNA
	269	50	1.6	71577	8	AC135587 Homo sapi
c	270	50	1.6	72402	8	AL389924 AL389924 Human DNA
	271	50	1.6	73656	14	AC016512 AC016512 Homo sapi
c	272	50	1.6	77001	8	AL513013 AL513013 Human DNA
c	273	50	1.6	79666	8	AL136969 AL136969 Human DNA
c	274	50	1.6	80427	8	AC010492 AC010492 Homo sapi
c	275	50	1.6	81490	8	AL590987 AL590987 Human DNA
c	276	50	1.6	82901	8	AL391335 AL391335 Human DNA
c	277	50	1.6	84031	8	AP000223 AP000223 Homo sapi
c	278	50	1.6	86800	8	AC055715 AC055715 Homo sapi
	279	50	1.6	88698	14	AC009008 AC009008 Homo sapi
c	280	50	1.6	90399	14	AL162405 AL162405 Homo sapi
	281	50	1.6	92487	8	AC068280 AC068280 Homo sapi
c	282	50	1.6	92872	8	AC022425 AC022425 Homo sapi
c	283	50	1.6	93381	14	HSDJ766D4 AL118504 Homo sapi
c	284	50	1.6	95097	8	AC022202 AC022202 Homo sapi
c	285	50	1.6	96299	8	AL133519 AL133519 Human DNA
c	286	50	1.6	96489	8	AC011395 AC011395 Homo sapi
	287	50	1.6	97687	8	AL591403 AL591403 Human DNA
c	288	50	1.6	98078	8	AL451083 AL451083 Human DNA
c	289	50	1.6	98842	8	AC008972 AC008972 Homo sapi
	290	50	1.6	99505	8	AP002087 AP002087 Homo sapi
c	291	50	1.6	100000	8	AP000085 AP000085 Homo sapi
c	292	50	1.6	100000	8	AP000137 AP000137 Homo sapi
	293	50	1.6	107278	8	AC114962 AC114962 Homo sapi
c	294	50	1.6	107714	14	AL356306 AL356306 Homo sapi
c	295	50	1.6	108162	14	AC020960 AC020960 Mus muscu
	296	50	1.6	110000	14	AC008576_1 Continuation (2 of
c	297	50	1.6	110000	14	AC079517_2 Continuation (3 of
	298	50	1.6	110000	14	AL590046_2 Continuation (3 of
	299	50	1.6	110753	8	AC010216 AC010216 Homo sapi
c	300	50	1.6	112707	8	AL359207 AL359207 Human DNA
	301	50	1.6	112954	14	AF302687 AF302687 Homo sapi
c	302	50	1.6	114148	8	AC112655 AC112655 Homo sapi
	303	50	1.6	114191	8	AL451007 AL451007 Human DNA
	304	50	1.6	114540	8	AC107426 AC107426 Homo sapi
c	305	50	1.6	114983	14	AC093209 AC093209 Homo sapi
	306	50	1.6	117520	8	AP005624 AP005624 Homo sapi
c	307	50	1.6	120129	8	AL391560 AL391560 Human DNA
	308	50	1.6	121282	8	AC027318 AC027318 Homo sapi
	309	50	1.6	123369	8	AC025170 AC025170 Homo sapi
	310	50	1.6	124102	8	AC021148 AC021148 Homo sapi
	311	50	1.6	125465	8	HSJ136J15 AL118496 Human DNA
	312	50	1.6	128450	8	AC108861 AC108861 Homo sapi
c	313	50	1.6	129272	8	AL451125 AL451125 Human DNA
	314	50	1.6	129800	14	AL163539 AL163539 Homo sapi
c	315	50	1.6	130336	8	HS17K7 AL035078 Human DNA
c	316	50	1.6	130454	14	AL391537 AL391537 Homo sapi
c	317	50	1.6	130647	8	AL157385 AL157385 Human DNA
	318	50	1.6	137206	14	AC024123 AC024123 Homo sapi
	319	50	1.6	138443	8	HS179E13 AL035603 Human DNA
c	320	50	1.6	139111	8	AL158048 AL158048 Human DNA
	321	50	1.6	139846	8	AC020895 AC020895 Homo sapi
c	322	50	1.6	140091	8	AL445684 AL445684 Human DNA
c	323	50	1.6	140195	8	AC087163 AC087163 Homo sapi
	324	50	1.6	140545	8	AC010719 AC010719 Homo sapi
	325	50	1.6	141670	8	AC104170 AC104170 Homo sapi
	326	50	1.6	141899	8	AC087457 AC087457 Homo sapi
	327	50	1.6	142225	8	AC110053 AC110053 Homo sapi
c	328	50	1.6	145340	8	AL136525 AL136525 Human DNA
c	329	50	1.6	145770	14	AC019244 AC019244 Homo sapi
c	330	50	1.6	146466	8	AL353637 AL353637 Human DNA
c	331	50	1.6	149103	8	AC013434 AC013434 Homo sapi
	332	50	1.6	149559	8	AL139327 AL139327 Human DNA
c	333	50	1.6	150228	8	AC003071 AC003071 Homo sapi
	334	50	1.6	150708	8	AP003467 AP003467 Homo sapi
c	335	50	1.6	151479	14	AC012278 AC012278 Homo sapi
c	336	50	1.6	151660	14	AC027450 AC027450 Homo sapi
c	337	50	1.6	152210	8	AY445095 AY445095 Homo sapi
	338	50	1.6	152846	14	AC025709 AC025709 Homo sapi
c	339	50	1.6	152905	14	AC044801 AC044801 Homo sapi
c	340	50	1.6	154567	14	AC151375 AC151375 Callithri
c	341	50	1.6	154592	8	AL591215 AL591215 Human DNA
c	342	50	1.6	155067	8	AC010301 AC010301 Homo sapi
	343	50	1.6	156094	8	AC090198 AC090198 Homo sapi
	344	50	1.6	157207	8	BX640542 BX640542 Human DNA
	345	50	1.6	157298	8	BX005266 BX005266 Human DNA
	346	50	1.6	157872	8	AC007850 AC007850 Homo sapi
c	347	50	1.6	157921	14	AC027726 AC027726 Homo sapi
	348	50	1.6	158255	14	AC025695 AC025695 Homo sapi
c	349	50	1.6	158302	14	AC069481 AC069481 Homo sapi
	350	50	1.6	158929	14	AC105089 AC105089 Homo sapi
c	351	50	1.6	159098	14	AC021980 AC021980 Homo sapi
c	352	50	1.6	160193	8	AC092442 AC092442 Homo sapi
c	353	50	1.6	161264	8	AC007011 AC007011 Homo sapi
	354	50	1.6	161306	14	AC091056 AC091056 Homo sapi
c	355	50	1.6	161507	8	AC046195 AC046195 Homo sapi
	356	50	1.6	161791	14	AC021690 AC021690 Homo sapi
	357	50	1.6	162100	8	AC018868 AC018868 Homo sapi
	358	50	1.6	163167	8	AC011347 AC011347 Homo sapi
	359	50	1.6	163520	8	AL391259 AL391259 Human DNA
c	360	50	1.6	163584	14	AC008470 AC008470 Homo sapi
c	361	50	1.6	163599	8	HSDJ81D8 AL109933 Human DNA
c	362	50	1.6	163954	8	AC090959 AC090959 Homo sapi
c	363	50	1.6	164118	14	AC021384 AC021384 Homo sapi
c	364	50	1.6	164236	8	BS000105 BS000105 Pan trogl
	365	50	1.6	164529	14	AC079975 AC079975 Homo sapi
c	366	50	1.6	164766	14	AC026657 AC026657 Homo sapi
	367	50	1.6	164774	14	AC019104 AC019104 Homo sapi
	368	50	1.6	164933	14	AC074231 AC074231 Homo sapi
c	369	50	1.6	165773	8	AL357037 AL357037 Human DNA
c	370	50	1.6	165942	8	AC104943 AC104943 Homo sapi
c	371	50	1.6	166687	8	BX284677 BX284677 Human DNA
	372	50	1.6	166876	14	AL356864 AL356864 Homo sapi
	373	50	1.6	166955	8	AC147030 AC147030 Pan trogl
c	374	50	1.6	167013	8	HSDJ828H9 AL121757 Human DNA
	375	50	1.6	167170	8	AC120782 AC120782 Pan trogl
	376	50	1.6	167293	14	AC022629 AC022629 Homo sapi
c	377	50	1.6	167386	14	AL159993 AL159993 Homo sapi
c	378	50	1.6	168111	8	HS525L6 AL023807 Human DNA
c	379	50	1.6	168114	8	AL354828 AL354828 Human DNA
	380	50	1.6	168235	8	AC104229 AC104229 Homo sapi
c	381	50	1.6	168493	8	AC013463 AC013463 Homo sapi
c	382	50	1.6	168499	8	AC092037 AC092037 Homo sapi
	383	50	1.6	168672	14	AC027102 AC027102 Homo sapi

	384	50	1.6	168753	8	AC006120 AC006120 Homo sapi
	385	50	1.6	168873	14	BX537112 BX537112 Homo sapi
c	386	50	1.6	168875	8	AC114794 AC114794 Homo sapi
c	387	50	1.6	168880	8	AC017104 AC017104 Homo sapi
	388	50	1.6	169166	8	AC027687 AC027687 Homo sapi
	389	50	1.6	169218	14	AC149013 AC149013 Papio anu
	390	50	1.6	169835	8	AP005717 AP005717 Homo sapi
c	391	50	1.6	170249	8	AC068799 AC068799 Homo sapi
c	392	50	1.6	171951	8	CNS05TDA AL356020 Human chr
c	393	50	1.6	172036	8	AC046181 AC046181 Homo sapi
c	394	50	1.6	172288	14	AL160164 AL160164 Homo sapi
	395	50	1.6	172514	14	AC146903 AC146903 Aotus nan
	396	50	1.6	172753	8	AL359259 AL359259 Human DNA
c	397	50	1.6	172816	8	AC093899 AC093899 Homo sapi
	398	50	1.6	172837	8	AL450304 AL450304 Human DNA
	399	50	1.6	173333	14	AC025282 AC025282 Homo sapi
	400	50	1.6	173706	14	AC022694 AC022694 Homo sapi
c	401	50	1.6	173952	14	AC080087 AC080087 Homo sapi
c	402	50	1.6	174023	8	AC080089 AC080089 Homo sapi
	403	50	1.6	174097	8	AC069513 AC069513 Homo sapi
c	404	50	1.6	174612	14	AC023889 AC023889 Homo sapi
	405	50	1.6	174693	8	AL356738 AL356738 Human DNA
c	406	50	1.6	174995	8	AL365502 AL365502 Human DNA
	407	50	1.6	175021	14	AC161086 AC161086 Papio ham
c	408	50	1.6	175347	14	AC009099 AC009099 Homo sapi
c	409	50	1.6	175440	14	AC137626 AC137626 Homo sapi
	410	50	1.6	175493	14	AC020769 AC020769 Homo sapi
	411	50	1.6	175822	14	AC068402 AC068402 Homo sapi
	412	50	1.6	176550	8	AL356356 AL356356 Human DNA
	413	50	1.6	176647	8	AC024163 AC024163 Homo sapi
	414	50	1.6	176671	14	AC080149 AC080149 Homo sapi
c	415	50	1.6	176708	14	AC027548 AC027548 Homo sapi
	416	50	1.6	176932	8	AC016772 AC016772 Homo sapi
	417	50	1.6	177448	8	AC044781 AC044781 Homo sapi
	418	50	1.6	177518	8	AC093577 AC093577 Homo sapi
c	419	50	1.6	177555	8	AC006450 AC006450 Homo sapi
	420	50	1.6	177688	8	CNS01DT5 AL132640 Human chr
c	421	50	1.6	177878	14	AC010893 AC010893 Homo sapi
c	422	50	1.6	178346	14	AC011885 AC011885 Homo sapi
c	423	50	1.6	179176	14	AC022857 AC022857 Homo sapi
	424	50	1.6	179327	8	AC149133 AC149133 Pan trogl
	425	50	1.6	181123	14	AC104064 AC104064 Homo sapi
c	426	50	1.6	181598	8	AC026172 AC026172 Homo sapi
c	427	50	1.6	181604	8	AC026194 AC026194 Homo sapi
c	428	50	1.6	181663	8	AC010894 AC010894 Homo sapi
	429	50	1.6	182567	14	AC145207 AC145207 Homo sapi
	430	50	1.6	182945	14	AC069004 AC069004 Homo sapi
	431	50	1.6	182996	8	CNS01DTU AL132986 Human chr
	432	50	1.6	183353	8	AC087071 AC087071 Homo sapi
	433	50	1.6	183539	8	AC092268 AC092268 Homo sapi
c	434	50	1.6	183818	14	AC149455 AC149455 Papio anu
	435	50	1.6	183842	8	AC130450 AC130450 Homo sapi
c	436	50	1.6	184059	14	AC024196 AC024196 Homo sapi
c	437	50	1.6	184536	14	AC068682 AC068682 Homo sapi
c	438	50	1.6	184706	8	AC120781 AC120781 Pan trogl
	439	50	1.6	184759	8	AC022695 AC022695 Homo sapi
	440	50	1.6	184906	8	CNS01DUQ AL133313 Human chr
c	441	50	1.6	186615	8	BX088717 BX088717 Human DNA
	442	50	1.6	187126	14	AP002395 AP002395 Homo sapi
	443	50	1.6	187239	14	AC021349 AC021349 Homo sapi
	444	50	1.6	187343	14	AC016688 AC016688 Homo sapi
	445	50	1.6	188636	8	HUAC002287 AC002287 Homo sapi
	446	50	1.6	190306	8	AC146080 AC146080 Pan trogl
c	447	50	1.6	191668	14	CR759969 CR759969 Homo sapi
c	448	50	1.6	192478	14	AC164657 AC164657 Pan trogl
	449	50	1.6	193882	8	AF313041 AF313041 Homo sapi
c	450	50	1.6	193889	14	AC151014 AC151014 Callithri
c	451	50	1.6	194919	14	AC127899 AC127899 Homo sapi
	452	50	1.6	195280	8	AC073869 AC073869 Homo sapi
	453	50	1.6	195383	14	AC126472 AC126472 Homo sapi
c	454	50	1.6	196096	14	AL161451 AL161451 Homo sapi
	455	50	1.6	196207	14	AC027048 AC027048 Homo sapi
c	456	50	1.6	196416	8	AC006483 AC006483 Homo sapi
	457	50	1.6	196990	14	AC156839 AC156839 Pan trogl
	458	50	1.6	197337	6	AR659697 AR659697 Sequence
	459	50	1.6	197819	8	AC152891 AC152891 Pan trogl
c	460	50	1.6	198017	8	AC012314 AC012314 Homo sapi
	461	50	1.6	198027	14	AC008515 AC008515 Homo sapi
	462	50	1.6	198253	8	AC009090 AC009090 Homo sapi
	463	50	1.6	198700	8	AC104317 AC104317 Homo sapi
	464	50	1.6	198720	8	AP002366 AP002366 Homo sapi
	465	50	1.6	198899	8	AC090192 AC090192 Homo sapi
	466	50	1.6	199218	8	AC108863 AC108863 Homo sapi
c	467	50	1.6	199566	8	AC103676 AC103676 Homo sapi
c	468	50	1.6	200000	14	AC004624 AC004624 Homo sapi
c	469	50	1.6	200000	14	AC004670 AC004670 Homo sapi
c	470	50	1.6	201414	8	AC090644 AC090644 Homo sapi
	471	50	1.6	202891	8	BX088651 BX088651 Human DNA
c	472	50	1.6	203838	14	AC026840 AC026840 Homo sapi
c	473	50	1.6	204102	14	AC013826 AC013826 Homo sapi
	474	50	1.6	204439	8	AP000807 AP000807 Homo sapi
c	475	50	1.6	204505	8	AC089983 AC089983 Homo sapi
c	476	50	1.6	205307	14	AC009968 AC009968 Homo sapi
	477	50	1.6	206258	14	AC112775 AC112775 Homo sapi
c	478	50	1.6	207815	14	AC053497 AC053497 Homo sapi
c	479	50	1.6	208117	8	AC008488 AC008488 Homo sapi
c	480	50	1.6	209532	14	AL356136 AL356136 Homo sapi
c	481	50	1.6	209888	14	AC146462 AC146462 Saimiri b
	482	50	1.6	210643	8	AC137723 AC137723 Homo sapi
c	483	50	1.6	210643	8	AC137723 AC137723 Homo sapi
c	484	50	1.6	211211	14	AC161277 AC161277 Pan trogl
c	485	50	1.6	211213	14	AC163770 AC163770 Pan trogl
	486	50	1.6	211550	8	AC073912 AC073912 Homo sapi
c	487	50	1.6	211791	8	AL162430 AL162430 Human DNA
	488	50	1.6	212622	8	AC025259 AC025259 Homo sapi
c	489	50	1.6	214978	8	AC007383 AC007383 Homo sapi
c	490	50	1.6	216129	14	AC157530 AC157530 Pan trogl
	491	50	1.6	217521	8	AC023825 AC023825 Homo sapi
	492	50	1.6	221887	8	AL158052 AL158052 Human DNA
c	493	50	1.6	226910	14	AC141314 AC141314 Homo sapi
c	494	50	1.6	233418	14	AC109128 AC109128 Homo sapi
	495	50	1.6	234287	6	AR659646 AR659646 Sequence
	496	50	1.6	234288	6	AR659800 AR659800 Sequence
	497	50	1.6	241178	14	AC120839 AC120839 Pan trogl
	498	50	1.6	243722	14	AC151855 AC151855 Papio ham
c	499	50	1.6	250529	8	HUAE000658 AE000658 Homo sapi
	500	50	1.6	295639	14	AL360019 AL360019 Homo sapi

ALIGNMENTS

```
RESULT 1
CQ841455
LOCUS       CQ841455               3122 bp    DNA     linear   PAT 02-AUG-2004
DEFINITION  Sequence 102 from Patent EP1440981.
ACCESSION   CQ841455
VERSION     CQ841455.1  GI:50893242
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.
  TITLE     Full-length human cdna
  JOURNAL   Patent: EP 1440981-A 102 28-JUL-2004;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .3122
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

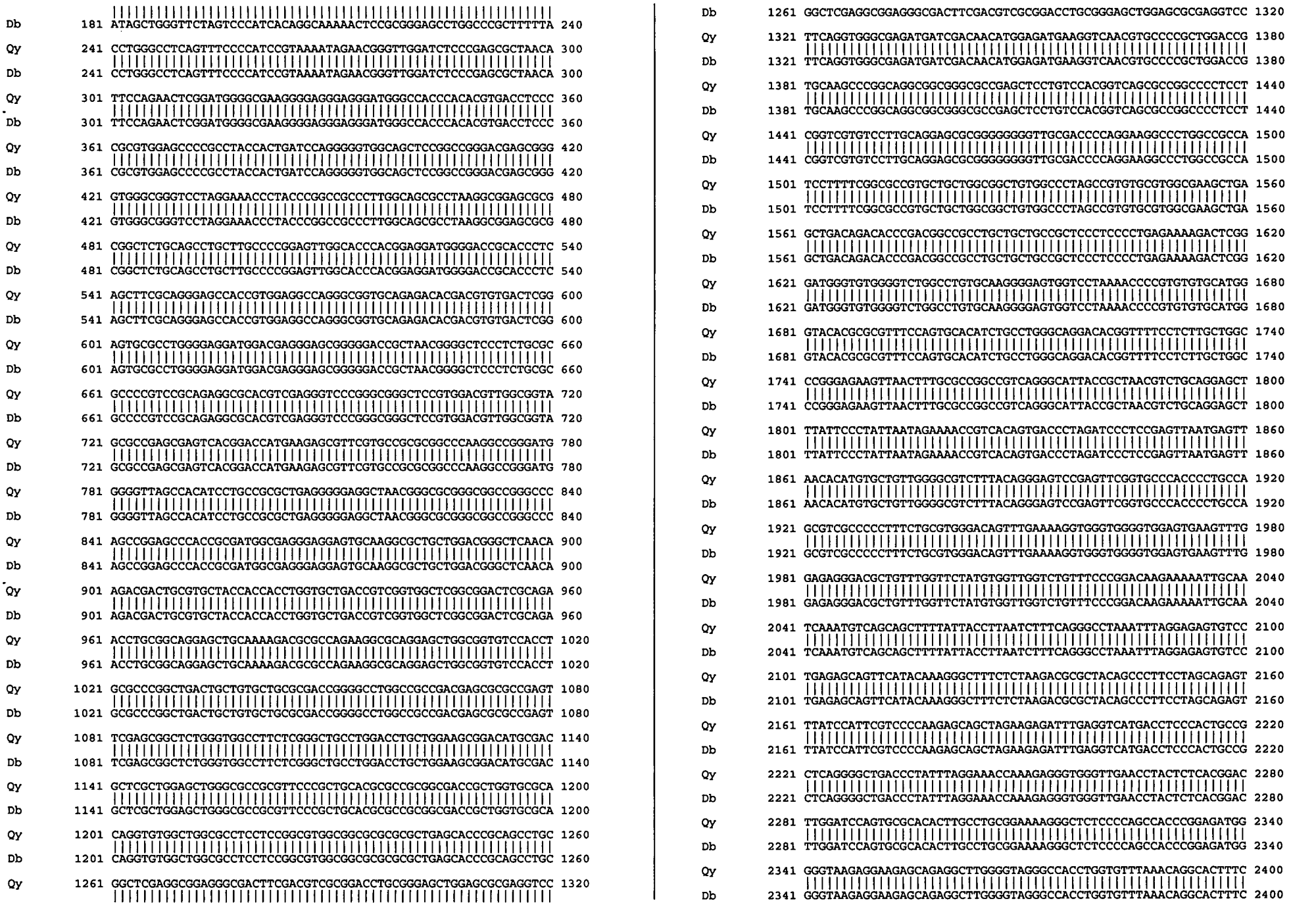
Query Match 100.0%; Score 3122; DB 6; Length 3122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440

Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920

Qy       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880

Qy       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940

Qy       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000

Qy       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060

Qy       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120

Qy       3121 TC 3122
                 ||
Db       3121 TC 3122
```

RESULT 2

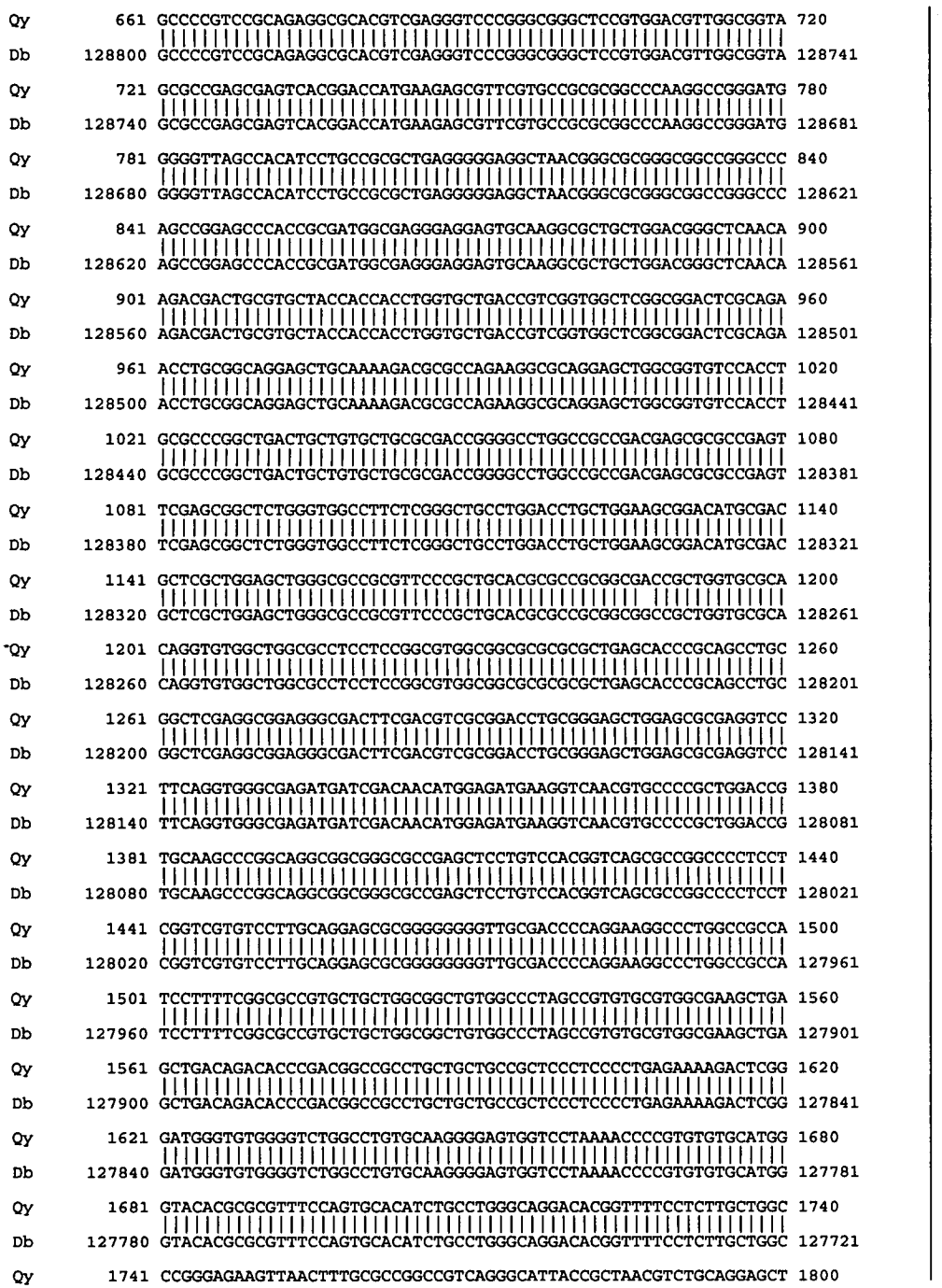
```
AK124499
LOCUS       AK124499               3122 bp    mRNA    linear   PRI 09-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ42508 fis, clone BRACE2043665.
ACCESSION   AK124499
VERSION     AK124499.1  GI:34530295
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3122)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .3122
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRACE2043665"
                     /tissue_type="cerebellum"
                     /clone_lib="BRACE2"
                     /note="cloning vector: pME18SFL3"
     CDS             858. .1565
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC85865.1"
                     /db_xref="GI:34530296"
                     /translation="MAREECKALLDGLNKTTACYHHLVLTVGGSADSQNLRQELQKTR
                     QKAQELAVSTCARLTAVLRDRGLAADERAEFERLWVAFSGCLDLLEADMRRSLELGAA
                     FPLHAPRRPLVRTGVAGASSGVAARALSTRSLRLEAEGDFDVADLRELEREVLQVGEM
                     IDNMEMKVNVPRWTVQARQAAGAELLSTVSAGPSSVVSLQERGGGCDPRKALAAILFG
                     AVLLAAVALAVCVAKLS"
ORIGIN

  Query Match           100.0%;  Score 3122;  DB 8;  Length 3122;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3122;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
```


Db 181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
Qy 241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
Db 241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
Qy 301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
Db 301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
Qy 361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
Db 361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
Qy 421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
Db 421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
Qy 481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
Db 481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
Qy 541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
Db 541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
Qy 601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
Db 601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
Qy 661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
Db 661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
Qy 721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
Db 721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
Qy 781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
Db 781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
Qy 841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
Db 841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
Qy 901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
Db 901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
Qy 961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Db 961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Qy 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Db 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Qy 1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Db 1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Qy 1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Db 1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Qy 1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Db 1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Qy 1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Db 1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Qy 1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Db 1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Qy 1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Db 1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Qy 1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Db 1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Qy 1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Db 1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Qy 1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Db 1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Qy 1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Db 1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Qy 1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Db 1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Qy 1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Db 1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Qy 1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Db 1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Qy 1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Db 1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Qy 1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Db 1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Qy 1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Db 1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Qy 2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Db 2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Qy 2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Db 2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Qy 2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Db 2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Qy 2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Db 2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Qy 2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Db 2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Qy 2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
Db 2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

```
Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880

Qy       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTTTTTTGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940

Qy       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGCGAAACCCGATCTCTACTAAAAAATATAAAAATTGGCCGGGCATGGTGGCGCATGCCT 3000

Qy       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GTGGTCCCAGCTACTCGGGAGGTTGAGGCAGGAGAGTCGCTTGAATGCAGGAGGTGGAGG 3060

Qy       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120

Qy       3121 TC 3122
              ||
Db       3121 TC 3122
```

```
RESULT 3
AC008736/c
LOCUS       AC008736              191925 bp    DNA     linear   PRI 27-SEP-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.
ACCESSION   AC008736
VERSION     AC008736.6  GI:10312244
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 191925)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 27, 2000 this sequence version replaced gi:8575905.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
            STS Content:
            SHGC-57769 G37408.
FEATURES             Location/Qualifiers
     source          1. .191925
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2538C1"
ORIGIN
```

```
  Query Match            91.0%;  Score 2840;  DB 8;  Length 191925;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2890;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129460 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 129401

Qy         61 ACCTGACCCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129400 ACCTGACCCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 129341

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129340 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 129281

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129280 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 129221

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129220 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 129161

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129160 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 129101

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129100 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 129041

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129040 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 128981

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128980 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 128921

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128920 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 128861

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128860 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 128801
```


Qy 661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
Db 128800 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 128741
Qy 721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
Db 128740 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 128681
Qy 781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
Db 128680 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 128621
Qy 841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
Db 128620 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 128561
Qy 901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
Db 128560 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 128501
Qy 961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
Db 128500 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 128441
Qy 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
Db 128440 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 128381
Qy 1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
Db 128380 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 128321
Qy 1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
Db 128320 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 128261
Qy 1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Db 128260 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 128201
Qy 1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Db 128200 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 128141
Qy 1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Db 128140 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 128081
Qy 1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Db 128080 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 128021
Qy 1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Db 128020 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 127961
Qy 1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Db 127960 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 127901
Qy 1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Db 127900 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 127841
Qy 1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Db 127840 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 127781
Qy 1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Db 127780 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 127721
Qy 1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Db 127720 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 127661
Qy 1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Db 127660 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 127601
Qy 1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Db 127600 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 127541
Qy 1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Db 127540 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 127481
Qy 1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Db 127480 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 127421
Qy 2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Db 127420 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 127361
Qy 2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Db 127360 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 127301
Qy 2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Db 127300 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 127241
Qy 2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Db 127240 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 127181
Qy 2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Db 127180 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 127121
Qy 2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
Db 127120 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 127061
Qy 2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
Db 127060 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 127001
Qy 2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
Db 127000 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 126941
Qy 2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
Db 126940 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 126881
Qy 2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
Db 126880 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 126821
Qy 2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
Db 126820 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 126761
Qy 2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
Db 126760 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 126701
Qy 2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
Db 126700 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 126641
Qy 2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 2880

```
Db      126640 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCACTTTTT 126581

Qy        2881 TCTTTTTTGAG 2891
               ||||||||||
Db      126580 TCTTTTTTGAG 126570


RESULT 4
AC027340/c
LOCUS       AC027340               216441 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT
            SEQUENCE, 52 unordered pieces.
ACCESSION   AC027340
VERSION     AC027340.2  GI:9211228
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 15, 2000 this sequence version replaced gi:7341654.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 726304, BC691328
            Center clone name: CITB-H1_2257C19
            -------
            Summary Statistics
            Consensus quality: 164680 bases at least Q40
            Consensus quality: 187929 bases at least Q30
            Consensus quality: 195340 bases at least Q20
            Estimated insert size: 159060; agarose-fp estimation
            Estimated insert size: 211341; sum-of-contigs estimation
            Quality coverage:  8.59 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.46 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1041: contig of 1041 bp in length
            *     1042     1141: gap of unknown length
            *     1142     2370: contig of 1229 bp in length
            *     2371     2470: gap of unknown length
            *     2471     3552: contig of 1082 bp in length
            *     3553     3652: gap of unknown length
            *     3653     4683: contig of 1031 bp in length
            *     4684     4783: gap of unknown length
            *     4784     6178: contig of 1395 bp in length
            *     6179     6278: gap of unknown length
            *     6279     7442: contig of 1164 bp in length
            *     7443     7542: gap of unknown length
            *     7543     8661: contig of 1119 bp in length
            *     8662     8761: gap of unknown length
            *     8762    10030: contig of 1269 bp in length
            *    10031    10130: gap of unknown length
            *    10131    11491: contig of 1361 bp in length
            *    11492    11591: gap of unknown length
            *    11592    12967: contig of 1376 bp in length
            *    12968    13067: gap of unknown length
            *    13068    14173: contig of 1106 bp in length
            *    14174    14273: gap of unknown length
            *    14274    15840: contig of 1567 bp in length
            *    15841    15940: gap of unknown length
            *    15941    17242: contig of 1302 bp in length
            *    17243    17342: gap of unknown length
            *    17343    18449: contig of 1107 bp in length
            *    18450    18549: gap of unknown length
            *    18550    19683: contig of 1134 bp in length
            *    19684    19783: gap of unknown length
            *    19784    21247: contig of 1464 bp in length
            *    21248    21347: gap of unknown length
            *    21348    22952: contig of 1605 bp in length
            *    22953    23052: gap of unknown length
            *    23053    24452: contig of 1400 bp in length
            *    24453    24552: gap of unknown length
            *    24553    25825: contig of 1273 bp in length
            *    25826    25925: gap of unknown length
            *    25926    27109: contig of 1184 bp in length
            *    27110    27209: gap of unknown length
            *    27210    28292: contig of 1083 bp in length
            *    28293    28392: gap of unknown length
            *    28393    29744: contig of 1352 bp in length
            *    29745    29844: gap of unknown length
            *    29845    31042: contig of 1198 bp in length
            *    31043    31142: gap of unknown length
            *    31143    32629: contig of 1487 bp in length
            *    32630    32729: gap of unknown length
            *    32730    33898: contig of 1169 bp in length
            *    33899    33998: gap of unknown length
            *    33999    35704: contig of 1706 bp in length
            *    35705    35804: gap of unknown length
            *    35805    37471: contig of 1667 bp in length
            *    37472    37571: gap of unknown length
            *    37572    39285: contig of 1714 bp in length
            *    39286    39385: gap of unknown length
            *    39386    40577: contig of 1192 bp in length
            *    40578    40677: gap of unknown length
            *    40678    41930: contig of 1253 bp in length
            *    41931    42030: gap of unknown length
            *    42031    43743: contig of 1713 bp in length
            *    43744    43843: gap of unknown length
            *    43844    46054: contig of 2211 bp in length
            *    46055    46154: gap of unknown length
            *    46155    49379: contig of 3225 bp in length
            *    49380    49479: gap of unknown length
            *    49480    53371: contig of 3892 bp in length
            *    53372    53471: gap of unknown length
            *    53472    55756: contig of 2285 bp in length
            *    55757    55856: gap of unknown length
            *    55857    58019: contig of 2163 bp in length
            *    58020    58119: gap of unknown length
            *    58120    60327: contig of 2208 bp in length
            *    60328    60427: gap of unknown length
            *    60428    64223: contig of 3796 bp in length
            *    64224    64323: gap of unknown length
            *    64324    67218: contig of 2895 bp in length
            *    67219    67318: gap of unknown length
            *    67319    71797: contig of 4479 bp in length
            *    71798    71897: gap of unknown length
            *    71898    76856: contig of 4959 bp in length
            *    76857    76956: gap of unknown length
            *    76957    82060: contig of 5104 bp in length
            *    82061    82160: gap of unknown length
            *    82161    87557: contig of 5397 bp in length
            *    87558    87657: gap of unknown length
            *    87658    94226: contig of 6569 bp in length
            *    94227    94326: gap of unknown length
            *    94327   101878: contig of 7552 bp in length
            *   101879   101978: gap of unknown length
            *   101979   110629: contig of 8651 bp in length
```

```
            *   110630   110729: gap of unknown length
            *   110730   119083: contig of 8354 bp in length
            *   119084   119183: gap of unknown length
            *   119184   128919: contig of 9736 bp in length
            *   128920   129019: gap of unknown length
            *   129020   142791: contig of 13772 bp in length
            *   142792   142891: gap of unknown length
            *   142892   156079: contig of 13188 bp in length
            *   156080   156179: gap of unknown length
            *   156180   183092: contig of 26913 bp in length
            *   183093   183192: gap of unknown length
            *   183193   216441: contig of 33249 bp in length.
FEATURES             Location/Qualifiers
     source          1. .216441
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2257C19"
                     /clone_lib="CalTech human BAC library D"
     gap             1042. .1141
                     /estimated_length=unknown
     gap             2371. .2470
                     /estimated_length=unknown
     gap             3553. .3652
                     /estimated_length=unknown
     gap             4684. .4783
                     /estimated_length=unknown
     gap             6179. .6278
                     /estimated_length=unknown
     gap             7443. .7542
                     /estimated_length=unknown
     gap             8662. .8761
                     /estimated_length=unknown
     gap             10031. .10130
                     /estimated_length=unknown
     gap             11492. .11591
                     /estimated_length=unknown
     gap             12968. .13067
                     /estimated_length=unknown
     gap             14174. .14273
                     /estimated_length=unknown
     gap             15841. .15940
                     /estimated_length=unknown
     gap             17243. .17342
                     /estimated_length=unknown
     gap             18450. .18549
                     /estimated_length=unknown
     gap             19684. .19783
                     /estimated_length=unknown
     gap             21248. .21347
                     /estimated_length=unknown
     gap             22953. .23052
                     /estimated_length=unknown
     gap             24453. .24552
                     /estimated_length=unknown
     gap             25826. .25925
                     /estimated_length=unknown
     gap             27110. .27209
                     /estimated_length=unknown
     gap             28293. .28392
                     /estimated_length=unknown
     gap             29745. .29844

 Query Match             91.0%;  Score 2840;  DB 14;  Length 216441;
 Best Local Similarity   99.9%;  Pred. No. 0;
 Matches 2890;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117034 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 116975

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116974 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 116915

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116914 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 116855

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116854 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 116795

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116794 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 116735

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116734 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 116675

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116674 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 116615

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116614 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 116555

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116554 CGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 116495

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116494 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 116435

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116434 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 116375

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116374 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 116315

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116314 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 116255

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116254 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 116195

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116194 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 116135

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116134 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 116075

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116074 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 116015

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116014 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 115955

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115954 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 115895

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
```

```
Db     115894 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 115835
Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
Db     115834 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 115775
Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
Db     115774 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 115715
Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
Db     115714 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 115655
Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
Db     115654 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 115595
Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
Db     115594 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 115535
Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
Db     115534 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 115475
Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
Db     115474 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 115415
Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
Db     115414 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 115355
Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
Db     115354 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 115295
Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
Db     115294 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 115235
Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860
Db     115234 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 115175
Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920
Db     115174 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 115115
Qy       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980
Db     115114 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 115055
Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040
Db     115054 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 114995
Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100
Db     114994 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 114935
Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160
Db     114934 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 114875
Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220
Db     114874 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 114815
Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280
Db     114814 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 114755
```

```
Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340
Db     114754 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 114695
Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400
Db     114694 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 114635
Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460
Db     114634 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 114575
Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520
Db     114574 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 114515
Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580
Db     114514 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 114455
Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640
Db     114454 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 114395
Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700
Db     114394 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 114335
Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760
Db     114334 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 114275
Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820
Db     114274 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 114215
Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 2880
Db     114214 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCCCACTTTTT 114155
Qy       2881 TCTTTTTTGAG 2891
Db     114154 TCTTTTTTGAG 114144
```

```
RESULT 5
AC008474
LOCUS       AC008474              186115 bp    DNA     linear   PRI 13-JUL-2002
DEFINITION  Homo sapiens chromosome 19 clone CTC-379B2, complete sequence.
ACCESSION   AC008474
VERSION     AC008474.9  GI:21743729
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 186115)
```

```
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 186115)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 13, 2002 this sequence version replaced gi:16197759.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.4.
FEATURES             Location/Qualifiers
     source          1. .186115
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTC-379B2"
ORIGIN

  Query Match             78.3%;  Score 2445;  DB 8;  Length 186115;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2865;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180657 ACTAGAGGTGGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGGCAGGGAC 180716

Qy         61 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180717 ACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCACTGCTCGCTTTGGGGCACG 180776

Qy        121 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180777 GAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGACAGGGTCCGGGAGTCAGT 180836

Qy        181 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180837 ATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTA 180896

Qy        241 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180897 CCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACA 180956

Qy        301 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180957 TTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCC 181016

Qy        361 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181017 CGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGG 181076

Qy        421 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181077 GTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCG 181136

Qy        481 CGGCTCTGCAGCCTGCTTGCCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181137 CGGCTCTGCAGCCTGCTTGCCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTC 181196

Qy        541 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181197 AGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGG 181256

Qy        601 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181257 AGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGC 181316

Qy        661 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181317 GCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTA 181376

Qy        721 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181377 GCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATG 181436

Qy        781 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181437 GGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCC 181496

Qy        841 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181497 AGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACA 181556

Qy        901 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181557 AGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGA 181616

Qy        961 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181617 ACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCT 181676

Qy       1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181677 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGT 181736

Qy       1081 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181737 TCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGAC 181796

Qy       1141 GCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCA 1200
              || |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     181797 GCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCA 181856

Qy       1201 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181857 CAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGC 181916

Qy       1261 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181917 GGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCC 181976

Qy       1321 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181977 TTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCG 182036

Qy       1381 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182037 TGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCT 182096

Qy       1441 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182097 CGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCA 182156

Qy       1501 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182157 TCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGA 182216

Qy       1561 GCTGACAGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182217 GCTGACGGACACCCGACGGCCGCCTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGG 182276

Qy       1621 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182277 GATGGGTGTGGGGTCTGGCCTGTGCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGG 182336

Qy       1681 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGC 1740
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     182337 GTACACGCGCGTTTCCAGTGCACATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGC 182395

Qy       1741 CCGGGAGAAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 1800
```

```
Db     182396 CCGGGAGGAGTTAACTTTGCGCCGGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCT 182455

Qy       1801 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 1860

Db     182456 TTATTCCCTATTAATAGAAAACCGTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTT 182515

Qy       1861 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 1920

Db     182516 AACACATGTGCTGTTGGGGCGTCTTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCA 182575

Qy       1921 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 1980

Db     182576 GCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTG 182635

Qy       1981 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 2040

Db     182636 GAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAA 182695

Qy       2041 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 2100

Db     182696 TCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCC 182755

Qy       2101 TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 2160

Db     182756 CGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGT 182815

Qy       2161 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 2220

Db     182816 TTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCG 182875

Qy       2221 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 2280

Db     182876 CTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGAC 182935

Qy       2281 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 2340

Db     182936 TTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGG 182995

Qy       2341 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 2400

Db     182996 GGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTC 183055

Qy       2401 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 2460

Db     183056 TCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCA 183115

Qy       2461 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 2520

Db     183116 GGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGC 183175

Qy       2521 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 2580

Db     183176 TGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGG 183235

Qy       2581 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 2640

Db     183236 GGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTA 183295

Qy       2641 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAA 2700

Db     183296 GAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATACTCTGAAA 183355

Qy       2701 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 2760

Db     183356 GAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGG 183415

Qy       2761 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 2820

Db     183416 GCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATACTTAGATT 183475

Qy       2821 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCC 2872

Db     183476 ACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCC 183527
```

RESULT 6
CQ850499

LOCUS	CQ850499 2327 bp DNA linear PAT 23-AUG-2004
DEFINITION	Sequence 968 from Patent EP1447413.
ACCESSION	CQ850499
VERSION	CQ850499.1 GI:51508711
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Isogai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T., Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
TITLE	Full-length human cDNA
JOURNAL	Patent: EP 1447413-A 968 18-AUG-2004; Research Association for Biotechnology (JP)
FEATURES	Location/Qualifiers
source	1. .2327 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	

Query Match 62.5%; Score 1952; DB 6; Length 2327;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503

Db          1 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 60

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563

Db         61 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 120

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623

Db        121 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 180

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683

Db        181 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 240

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743

Db        241 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 300

Qy        744 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 803

Db        301 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 360

Qy        804 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 863

Db        361 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 420

Qy        864 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 923

Db        421 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 480

Qy        924 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 983

Db        481 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 540

Qy        984 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 1043

Db        541 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 600

Qy       1044 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
```

```
Db        601 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 660

Qy       1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163

Db        661 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGCGCCGCG 720

Qy       1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223

Db        721 TTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 780

Qy       1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283

Db        781 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 840

Qy       1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343

Db        841 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 900

Qy       1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403

Db        901 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 960

Qy       1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1463

Db        961 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1020

Qy       1464 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1523

Db       1021 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1080

Qy       1524 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGC 1583

Db       1081 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACGGACACCCGACGGCCGC 1140

Qy       1584 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1643

Db       1141 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1200

Qy       1644 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1703

Db       1201 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1260

Qy       1704 ATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCC 1763

Db       1261 ATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGCCCGGGAGGAGTTAACTTTGCGCC 1319

Qy       1764 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1823

Db       1320 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1379

Qy       1824 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1883

Db       1380 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1439

Qy       1884 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1943

Db       1440 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1499

Qy       1944 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 2003

Db       1500 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 1559

Qy       2004 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 2063

Db       1560 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 1619

Qy       2064 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGC 2123

Db       1620 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGC 1679

Qy       2124 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 2183

Db       1680 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 1739

Qy       2184 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 2243

Db       1740 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 1799

Qy       2244 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 2303

Db       1800 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 1859

Qy       2304 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 2363

Db       1860 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 1919

Qy       2364 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 2423

Db       1920 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 1979

Qy       2424 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2483

Db       1980 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2039

Qy       2484 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2543

Db       2040 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2099

Qy       2544 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2603

Db       2100 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2159

Qy       2604 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2663

Db       2160 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2219

Qy       2664 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2723

Db       2220 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2279

Qy       2724 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2771

Db       2280 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2327
```

```
RESULT 7
AK127646
LOCUS       AK127646               2327 bp    mRNA    linear   PRI 19-FEB-2004
DEFINITION  Homo sapiens cDNA FLJ45744 fis, clone KIDNE2017153.
ACCESSION   AK127646
VERSION     AK127646.1  GI:34534649
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
            Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and
            Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2327)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
```

Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: Reverse Proteomics Research Institute, HRI and RAB.

```
FEATURES             Location/Qualifiers
     source          1. .2327
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KIDNE2017153"
                     /tissue_type="kidney"
                     /clone_lib="KIDNE2"
                     /note="cloning vector: pME18SFL3"
     CDS             415. .1122
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC87072.1"
                     /db_xref="GI:34534650"
                     /translation="MAREECKALLDGLNKTTACYHHLVLTVGGSADSQNLRQELQKTR
                     QKAQELAVSTCARLTAVLRDRGLAADERAEFERLWVAFSGCLDLLEADMRRALELGAA
                     FPLHAPRRPLVRTGVAGASSGVAARALSTRSLRLEAEGDFDVADLRELEREVLQVGEM
                     IDNMEMKVNVPRWTVQARQAAGAELLSTVSAGPSSVVSLQERGGGCDPRKALAAILFG
                     AVLLAAVALAVCVAKLS"
ORIGIN
```

Query Match 62.5%; Score 1952; DB 8; Length 2327;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
Db          1 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 60

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
Db         61 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 120

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
Db        121 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 180

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
Db        181 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 240

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
Db        241 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 300

Qy        744 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 803
Db        301 GAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCG 360

Qy        804 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 863
Db        361 CTGAGGGGGAGGCTAACGGGCGCGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCG 420

Qy        864 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 923
Db        421 AGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTACCACCAC 480

Qy        924 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 983
Db        481 CTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAG 540

Qy        984 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 1043
Db        541 ACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTG 600

Qy       1044 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 1103
Db        601 CGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTC 660

Qy       1104 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCG 1163
Db        661 TCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGCGCCGCG 720

Qy       1164 TTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 1223
Db        721 TTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCCTCCTCC 780

Qy       1224 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 1283
Db        781 GGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGCGACTTC 840

Qy       1284 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 1343
Db        841 GACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATGATCGAC 900

Qy       1344 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 1403
Db        901 AACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCGGCGGGC 960

Qy       1404 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1463
Db        961 GCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAGGAGCGC 1020

Qy       1464 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1523
Db       1021 GGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTGCTGCTG 1080

Qy       1524 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACAGACACCCGACGGCCGC 1583
Db       1081 GCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGACGGACACCCGACGGCCGC 1140

Qy       1584 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1643
Db       1141 CTGCTGCTGCCGCTCCCTCCCCTGAGAAAAGACTCGGGATGGGTGTGGGGTCTGGCCTGT 1200

Qy       1644 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1703
Db       1201 GCAAGGGGAGTGGTCCTAAAACCCCGTGTGTGCATGGGTACACGCGCGTTTCCAGTGCAC 1260

Qy       1704 ATCTGCCTGGGCAGGACACGGTTTTCCTCTTGCTGGCCCGGGAGAAGTTAACTTTGCGCC 1763
Db       1261 ATCTGCCTGGGCAGGACACGG-TTTCCTCTTGCTGGCCCGGGAGGAGTTAACTTTGCGCC 1319

Qy       1764 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1823
Db       1320 GGCCGTCAGGGCATTACCGCTAACGTCTGCAGGAGCTTTATTCCCTATTAATAGAAAACC 1379

Qy       1824 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1883
Db       1380 GTCACAGTGACCCTAGATCCCTCCGAGTTAATGAGTTAACACATGTGCTGTTGGGGCGTC 1439

Qy       1884 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1943
Db       1440 TTTACAGGGAGTCCGAGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGG 1499

Qy       1944 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 2003
Db       1500 ACAGTTTGAAAAGGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTA 1559

Qy       2004 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 2063
Db       1560 TGTGGTTGGTCTGTTTCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTA 1619

Qy       2064 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGC 2123
Db       1620 CCTTAATCTTTCAGGGCCTAAATTTAGGAGAGTGTCCCGAGAGCAGTTCATACAAAGGGC 1679

Qy       2124 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 2183
```

```
Db       1680 TTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCA 1739

Qy       2184 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 2243

Db       1740 GCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGG 1799

Qy       2244 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 2303

Db       1800 AAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGC 1859

Qy       2304 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 2363

Db       1860 CTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCT 1919

Qy       2364 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 2423

Db       1920 TGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTG 1979

Qy       2424 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2483

Db       1980 TTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAG 2039

Qy       2484 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2543

Db       2040 AGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTT 2099

Qy       2544 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2603

Db       2100 GGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCA 2159

Qy       2604 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2663

Db       2160 CAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGG 2219

Qy       2664 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2723

Db       2220 ATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGT 2279

Qy       2724 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2771

Db       2280 GTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTC 2327
```

```
RESULT 8
CQ748891
LOCUS       CQ748891                708 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34825 from Patent WO02068579.
ACCESSION   CQ748891
VERSION     CQ748891.1  GI:42376058
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 34825 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             21.0%;  Score 657;  DB 6;  Length 708;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  707;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917

Db          1 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 60

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977

Db         61 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 120

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037

Db        121 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 180

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097

Db        181 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 240

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157

Db        241 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 300

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217

Db        301 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 360

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277

Db        361 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 420

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337

Db        421 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 480

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397

Db        481 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 540

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457

Db        541 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 600

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517

Db        601 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 660

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565

Db        661 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 708
```

```
RESULT 9
HUMZD85E06
LOCUS       HUMZD85E06              655 bp    mRNA    linear   PRI 29-AUG-1998
DEFINITION  Homo sapiens full length insert cDNA clone ZD85E06.
ACCESSION   AF086461
VERSION     AF086461.1  GI:3483806
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Woessner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,
            Marth,G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theising,B.,
            Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
            Gibbons,M., Harvey,N., Pape,D., Chamberlain,A., Morales,R.,
            Schurk,R., Ritter,E., Kohn,S., Swaller,T., Behymer,K., Hillier,L.,
            Wilson,R. and Waterston,R.
  TITLE     Full Clone Sequencing of the Longest Available Member from Each
            Unigene Cluster
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 655)
```

```
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     SUBMITTED BY:
            Genome Sequencing Center
            Department of Genetics
            Washington University
            St. Louis MO 63108, USA
            http://genome.wustl.edu/gsc
            mailto:est@watson.wustl.edu

            NOTICE:  This sequence represents the full insert of this cDNA.  No
            attempt has been made to verify whether this corresponds to the
            full-length of the original mRNA from which it was derived.  We
            have tried to obtain double-stranded, or double chemistry sequence
            across the entire clone, but potentially, there are areas in the
            sequence where this level of coverage was not achieved.
            Nevertheless, we are confident of the accuracy of this sequence as
            all regions of low quality, as defined by PHRAP (P. Green, in
            preparation), were visually inspected and edited accordingly.  The
            consensus quality values for this sequence have been submitted
            separately.

            The location of this clone is unknown.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:347458"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
ORIGIN

  Query Match            20.7%;  Score 646;  DB 8;  Length 655;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  646;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2129 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 60

Qy       2189 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 2248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 120

Qy       2249 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 180

Qy       2309 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 2368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 240

Qy       2369 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 2428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 300

Qy       2429 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 2488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 360

Qy       2489 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 2548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 420

Qy       2549 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 2608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 480

Qy       2609 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 2668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 540

Qy       2669 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 2728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 600

Qy       2729 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             |||||||||||||||||||||||||||||||||||||||||||||
Db          601 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 646


RESULT 10
CQ746021
LOCUS       CQ746021                 708 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31955 from Patent WO02068579.
ACCESSION   CQ746021
VERSION     CQ746021.1  GI:42365754
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 31955 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            19.4%;  Score 606;  DB 6;  Length 708;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches  706;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 60

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 120

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 180

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 240

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         241 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCGCGCTGGAGCTGGGC 300

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         301 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 360

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 420

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 480

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 540

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 600

Qy       1458 GAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGCGCGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 660

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 708


RESULT 11
AC027340
LOCUS       AC027340               216441 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT
            SEQUENCE, 52 unordered pieces.
ACCESSION   AC027340
VERSION     AC027340.2  GI:9211228
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216441)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 15, 2000 this sequence version replaced gi:7341654.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 726304, BC691328
            Center clone name: CITB-H1_2257C19
            -------
            Summary Statistics
            Consensus quality: 164680 bases at least Q40
            Consensus quality: 187929 bases at least Q30
            Consensus quality: 195340 bases at least Q20
            Estimated insert size: 159060; agarose-fp estimation
            Estimated insert size: 211341; sum-of-contigs estimation
            Quality coverage:  8.59 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.46 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1041: contig of 1041 bp in length
            *     1042     1141: gap of unknown length
            *     1142     2370: contig of 1229 bp in length
            *     2371     2470: gap of unknown length
            *     2471     3552: contig of 1082 bp in length
            *     3553     3652: gap of unknown length
            *     3653     4683: contig of 1031 bp in length
            *     4684     4783: gap of unknown length
            *     4784     6178: contig of 1395 bp in length
            *     6179     6278: gap of unknown length
            *     6279     7442: contig of 1164 bp in length
            *     7443     7542: gap of unknown length
            *     7543     8661: contig of 1119 bp in length
            *     8662     8761: gap of unknown length
            *     8762    10030: contig of 1269 bp in length
            *    10031    10130: gap of unknown length
            *    10131    11491: contig of 1361 bp in length
            *    11492    11591: gap of unknown length
            *    11592    12967: contig of 1376 bp in length
            *    12968    13067: gap of unknown length
            *    13068    14173: contig of 1106 bp in length
            *    14174    14273: gap of unknown length
            *    14274    15840: contig of 1567 bp in length
            *    15841    15940: gap of unknown length
            *    15941    17242: contig of 1302 bp in length
            *    17243    17342: gap of unknown length
            *    17343    18449: contig of 1107 bp in length
            *    18450    18549: gap of unknown length
            *    18550    19683: contig of 1134 bp in length
            *    19684    19783: gap of unknown length
            *    19784    21247: contig of 1464 bp in length
            *    21248    21347: gap of unknown length
            *    21348    22952: contig of 1605 bp in length
            *    22953    23052: gap of unknown length
            *    23053    24452: contig of 1400 bp in length
            *    24453    24552: gap of unknown length
            *    24553    25825: contig of 1273 bp in length
            *    25826    25925: gap of unknown length
            *    25926    27109: contig of 1184 bp in length
            *    27110    27209: gap of unknown length
            *    27210    28292: contig of 1083 bp in length
            *    28293    28392: gap of unknown length
            *    28393    29744: contig of 1352 bp in length
            *    29745    29844: gap of unknown length
            *    29845    31042: contig of 1198 bp in length
            *    31043    31142: gap of unknown length
            *    31143    32629: contig of 1487 bp in length
            *    32630    32729: gap of unknown length
            *    32730    33898: contig of 1169 bp in length
            *    33899    33998: gap of unknown length
            *    33999    35704: contig of 1706 bp in length
            *    35705    35804: gap of unknown length
            *    35805    37471: contig of 1667 bp in length
            *    37472    37571: gap of unknown length
            *    37572    39285: contig of 1714 bp in length
            *    39286    39385: gap of unknown length
            *    39386    40577: contig of 1192 bp in length
            *    40578    40677: gap of unknown length
            *    40678    41930: contig of 1253 bp in length
            *    41931    42030: gap of unknown length
            *    42031    43743: contig of 1713 bp in length
            *    43744    43843: gap of unknown length
            *    43844    46054: contig of 2211 bp in length
            *    46055    46154: gap of unknown length
            *    46155    49379: contig of 3225 bp in length
            *    49380    49479: gap of unknown length
            *    49480    53371: contig of 3892 bp in length
            *    53372    53471: gap of unknown length
            *    53472    55756: contig of 2285 bp in length
            *    55757    55856: gap of unknown length
            *    55857    58019: contig of 2163 bp in length
            *    58020    58119: gap of unknown length
            *    58120    60327: contig of 2208 bp in length
            *    60328    60427: gap of unknown length
            *    60428    64223: contig of 3796 bp in length
            *    64224    64323: gap of unknown length
            *    64324    67218: contig of 2895 bp in length
            *    67219    67318: gap of unknown length
            *    67319    71797: contig of 4479 bp in length
            *    71798    71897: gap of unknown length
            *    71898    76856: contig of 4959 bp in length
            *    76857    76956: gap of unknown length
```

```
            *    76957    82060: contig of 5104 bp in length
            *    82061    82160: gap of unknown length
            *    82161    87557: contig of 5397 bp in length
            *    87558    87657: gap of unknown length
            *    87658    94226: contig of 6569 bp in length
            *    94227    94326: gap of unknown length
            *    94327   101878: contig of 7552 bp in length
            *   101879   101978: gap of unknown length
            *   101979   110629: contig of 8651 bp in length
            *   110630   110729: gap of unknown length
            *   110730   119083: contig of 8354 bp in length
            *   119084   119183: gap of unknown length
            *   119184   128919: contig of 9736 bp in length
            *   128920   129019: gap of unknown length
            *   129020   142791: contig of 13772 bp in length
            *   142792   142891: gap of unknown length
            *   142892   156079: contig of 13188 bp in length
            *   156080   156179: gap of unknown length
            *   156180   183092: contig of 26913 bp in length
            *   183093   183192: gap of unknown length
            *   183193   216441: contig of 33249 bp in length.
FEATURES             Location/Qualifiers
     source          1. .216441
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2257C19"
                     /clone_lib="CalTech human BAC library D"
     gap             1042. .1141
                     /estimated_length=unknown
     gap             2371. .2470
                     /estimated_length=unknown
     gap             3553. .3652
                     /estimated_length=unknown
     gap             4684. .4783
                     /estimated_length=unknown
     gap             6179. .6278
                     /estimated_length=unknown
     gap             7443. .7542
                     /estimated_length=unknown
     gap             8662. .8761
                     /estimated_length=unknown
     gap             10031. .10130
                     /estimated_length=unknown
     gap             11492. .11591
                     /estimated_length=unknown
     gap             12968. .13067
                     /estimated_length=unknown
     gap             14174. .14273
                     /estimated_length=unknown
     gap             15841. .15940
                     /estimated_length=unknown
     gap             17243. .17342
                     /estimated_length=unknown
     gap             18450. .18549
                     /estimated_length=unknown
     gap             19684. .19783
                     /estimated_length=unknown
     gap             21248. .21347
                     /estimated_length=unknown
     gap             22953. .23052
                     /estimated_length=unknown
     gap             24453. .24552
                     /estimated_length=unknown
     gap             25826. .25925
                     /estimated_length=unknown
     gap             27110. .27209
                     /estimated_length=unknown
     gap             28293. .28392
                     /estimated_length=unknown
     gap             29745. .29844
```

```
  Query Match           12.6%;  Score 392;  DB 14;  Length 216441;
  Best Local Similarity  99.6%;  Pred. No. 3.5e-202;
  Matches  752;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    3;

Qy       2036 TGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAG 2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42288 TGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTAAATTTAGGAGAG 42347

Qy       2096 TGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGC 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42348 TGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGC 42407

Qy       2156 AGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTCATGACCTCCCAC 2215
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      42408 AGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGA-TTGAGGTCATGACCTCCCAC 42466

Qy       2216 TGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCA 2275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42467 TGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCA 42526

Qy       2276 CGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGA 2335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42527 CGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGA 42586

Qy       2336 GATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCA 2395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42587 GATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCA 42646

Qy       2396 CTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCT 2455
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      42647 CTTTCTCCTTCTCTGGGGCTTA-TTTTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCT 42705

Qy       2456 TTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTG 2515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42706 TTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTG 42765

Qy       2516 CGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCC 2575
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      42766 CGTGCTGGCCAGTAGGAGGGCTGGC-TTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCC 42824

Qy       2576 TCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTG 2635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42825 TCCGGGGTAATGTGGCATTACTGGCCCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTG 42884

Qy       2636 GGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTC 2695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42885 GGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTC 42944

Qy       2696 TGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAAC 2755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42945 TGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAGAAC 43004

Qy       2756 GCTGGGCCATGAACTCATACCTGCCAATGAGTCAA 2790
             |||||||||||||||||||||||||||||||||||
Db      43005 GCTGGGCCATGAACTCATACCTGCCAATGAGTCAA 43039
```

```
RESULT 12
G37408/c
LOCUS       G37408                   449 bp    DNA     linear   STS 31-MAR-1998
DEFINITION  SHGC-57769 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION   G37408
VERSION     G37408.1  GI:2997059
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Myers,R.M.
  TITLE     Human STSs (1997)
  JOURNAL   Unpublished (1997)
```

```
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: CATGGCCCAGCGTTCTTTAT
            Primer B: ACTGGCCCACAGAGGTTTTG
            STS size: 173
            PCR Profile:
                    Initial incubation: 95 degrees C for 10 minutes

                    Denaturation:       94 degrees C for 30 seconds
                    Annealing:          60 degrees C for 30 seconds
                    Polymerization:     72 degrees C for 23 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9700
            Protocol:
                    Template:                   25 ng
                    Primer:                     each 1 uM
                    dNTPs:                      each 200 uM
                    AmpliTaq Gold Polymerase:   0.07 units/ul
                    Total Vol:                  5 ul

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:            50 mM
                    Tris-HCl:       10 mM
                    pH:             8.3

            Prepared with primer pairs derived from W81222 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="19"
                     /clone_lib="Human"
     STS             10. .182
     primer_bind     10. .29
     primer_bind     complement(163. .182)
ORIGIN

  Query Match             4.6%;  Score 144;  DB 10;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-66;
  Matches  144;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2631 GACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAA 2690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAA 86

Qy       2691 TCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAA 2750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 TCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAA 26

Qy       2751 AGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||
Db          25 AGAACGCTGGGCCATGAACTCATA 2


RESULT 13
CQ553695
LOCUS       CQ553695                  60 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 23330 from Patent WO0210449.
ACCESSION   CQ553695
VERSION     CQ553695.1  GI:41520122
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
  TITLE     Oligonucleotide library for detecting rna transcripts and splice
            variants that populate a transcriptome
  JOURNAL   Patent: WO 0210449-A 23330 07-FEB-2002;
            Compugen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             1.9%;  Score 60;  DB 6;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2464 GGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGG 2523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGG 60


RESULT 14
CR762434
LOCUS       CR762434                24495 bp    DNA     linear   PRI 16-FEB-2005
DEFINITION  Human DNA sequence from clone DADB-91M20 on chromosome 6, complete
            sequence.
ACCESSION   CR762434
VERSION     CR762434.7  GI:59891252
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 24495)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 16, 2005 this sequence version replaced gi:56369811.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            DADB-91M20 is from a DNA-arts DBB human bac library VECTOR:
```

```
          pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .24495
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DADB-91M20"
                     /clone_lib="DNA-arts-BAC.1-DBB.1"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 24495;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8493 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8545


RESULT 15
D84401/c
LOCUS       D84401                42398 bp    DNA     linear   PRI 20-NOV-1999
DEFINITION  Homo sapiens genomic DNA, 43 kb segment from chromosome6, complete
            sequence.
ACCESSION   D84401
VERSION     D84401.1  GI:2401263
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kikuti,Y.Y., Tamiya,G., Ando,A., Chen,L., Kimura,M., Ferreira,E.,
            Tsuji,K., Trowsdale,J. and Inoko,H.
  TITLE     Physical mapping 220 kb centromeric of the human MHC and DNA
            sequence analysis of the 43-kb segment including the RING1, HKE6,
            and HKE4 genes
  JOURNAL   Genomics 42 (3), 422-435 (1997)
   PUBMED   9205114
REFERENCE   2  (bases 1 to 42398)
  AUTHORS   Kikuti,Y., Inoko,H., Ando,A., Kimura,M., Watanabe,K. and Shiina,T.
  TITLE     Physical map of 200 kb at the centromeric side of the human MHC
            region and sequence of 43 kb
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 42398)
  AUTHORS   Inoko,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-1996) Hidetoshi Inoko, Tokai University School of
            Medicine, Molecular Life Science; Bohseidai, Isehara, Kanagawa
            259-11, Japan (E-mail:hinoko@is.icc.u-tokai.ac.jp,
            Tel:0463-93-1121, Fax:0463-94-8884)
FEATURES             Location/Qualifiers
     source          1. .42398
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="519"
                     /cell_line="CGMI"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 42398;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5305 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 5253


RESULT 16
CR547129
LOCUS       CR547129               49878 bp    DNA     linear   PRI 18-JAN-2005
DEFINITION  Human DNA sequence from clone DAMA-258A19 on chromosome 6, complete
            sequence.
ACCESSION   CR547129
VERSION     CR547129.3  GI:57898954
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 49878)
  AUTHORS   Wood,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 18, 2005 this sequence version replaced gi:50299847.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAMA-258A19 is from the DNA-Arts human BAC library MANN.1 VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .49878
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMA-258A19"
                     /clone_lib="DNA-Arts BAC library MANN.1"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 49878;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30123 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 30175


RESULT 17
AC061985
LOCUS       AC061985               62449 bp    DNA     linear   HTG 21-APR-2000
DEFINITION  Homo sapiens chromosome 17 clone -2009I15 map 17, LOW-PASS SEQUENCE
```

```
            SAMPLING.
ACCESSION   AC061985
VERSION     AC061985.1  GI:7630724
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 62449)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone -2009I15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 62449)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9672
                Center clone name: 2009_I_15
            --------------.
            * NOTE: This record contains 72 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      753: contig of 753 bp in length
            *      754      853: gap of 100 bp
            *      854     1619: contig of 766 bp in length
            *     1620     1719: gap of 100 bp
            *     1720     2489: contig of 770 bp in length
            *     2490     2589: gap of 100 bp
            *     2590     3367: contig of 778 bp in length
            *     3368     3467: gap of 100 bp
            *     3468     4239: contig of 772 bp in length
            *     4240     4339: gap of 100 bp
            *     4340     5105: contig of 766 bp in length
            *     5106     5205: gap of 100 bp
            *     5206     5967: contig of 762 bp in length
            *     5968     6067: gap of 100 bp
            *     6068     6846: contig of 779 bp in length
            *     6847     6946: gap of 100 bp
            *     6947     7717: contig of 771 bp in length
            *     7718     7817: gap of 100 bp
            *     7818     8586: contig of 769 bp in length
            *     8587     8686: gap of 100 bp
            *     8687     9488: contig of 802 bp in length
            *     9489     9588: gap of 100 bp
            *     9589    10353: contig of 765 bp in length
            *    10354    10453: gap of 100 bp
            *    10454    11229: contig of 776 bp in length
            *    11230    11329: gap of 100 bp
            *    11330    12086: contig of 757 bp in length
            *    12087    12186: gap of 100 bp
            *    12187    12958: contig of 772 bp in length
            *    12959    13058: gap of 100 bp
            *    13059    13841: contig of 783 bp in length
            *    13842    13941: gap of 100 bp
            *    13942    14700: contig of 759 bp in length
            *    14701    14800: gap of 100 bp
            *    14801    15557: contig of 757 bp in length
            *    15558    15657: gap of 100 bp
            *    15658    16431: contig of 774 bp in length
            *    16432    16531: gap of 100 bp
            *    16532    17306: contig of 775 bp in length
            *    17307    17406: gap of 100 bp
            *    17407    18178: contig of 772 bp in length
            *    18179    18278: gap of 100 bp
            *    18279    19032: contig of 754 bp in length
            *    19033    19132: gap of 100 bp
            *    19133    19883: contig of 751 bp in length
            *    19884    19983: gap of 100 bp
            *    19984    20761: contig of 778 bp in length
            *    20762    20861: gap of 100 bp
            *    20862    21636: contig of 775 bp in length
            *    21637    21736: gap of 100 bp
            *    21737    22520: contig of 784 bp in length
            *    22521    22620: gap of 100 bp
            *    22621    23398: contig of 778 bp in length
            *    23399    23498: gap of 100 bp
            *    23499    24277: contig of 779 bp in length
            *    24278    24377: gap of 100 bp
            *    24378    25163: contig of 786 bp in length
            *    25164    25263: gap of 100 bp
            *    25264    26026: contig of 763 bp in length
            *    26027    26126: gap of 100 bp
            *    26127    26892: contig of 766 bp in length
            *    26893    26992: gap of 100 bp
            *    26993    27753: contig of 761 bp in length
            *    27754    27853: gap of 100 bp
            *    27854    28621: contig of 768 bp in length
            *    28622    28721: gap of 100 bp
            *    28722    29499: contig of 778 bp in length
            *    29500    29599: gap of 100 bp
            *    29600    30370: contig of 771 bp in length
            *    30371    30470: gap of 100 bp
            *    30471    31247: contig of 777 bp in length
            *    31248    31347: gap of 100 bp
            *    31348    32118: contig of 771 bp in length
            *    32119    32218: gap of 100 bp
            *    32219    32946: contig of 728 bp in length
            *    32947    33046: gap of 100 bp
            *    33047    33824: contig of 778 bp in length
            *    33825    33924: gap of 100 bp
            *    33925    34678: contig of 754 bp in length
            *    34679    34778: gap of 100 bp
            *    34779    35551: contig of 773 bp in length
            *    35552    35651: gap of 100 bp
            *    35652    36422: contig of 771 bp in length
            *    36423    36522: gap of 100 bp
            *    36523    37295: contig of 773 bp in length
```

```
            *    37296    37395: gap of 100 bp
            *    37396    38160: contig of 765 bp in length
            *    38161    38260: gap of 100 bp
            *    38261    39025: contig of 765 bp in length
            *    39026    39125: gap of 100 bp
            *    39126    39879: contig of 754 bp in length
            *    39880    39979: gap of 100 bp
            *    39980    40746: contig of 767 bp in length
            *    40747    40846: gap of 100 bp
            *    40847    41620: contig of 774 bp in length
            *    41621    41720: gap of 100 bp
            *    41721    42483: contig of 763 bp in length
            *    42484    42583: gap of 100 bp
            *    42584    43359: contig of 776 bp in length
            *    43360    43459: gap of 100 bp
            *    43460    44236: contig of 777 bp in length
            *    44237    44336: gap of 100 bp
            *    44337    45111: contig of 775 bp in length
            *    45112    45211: gap of 100 bp
            *    45212    45981: contig of 770 bp in length
            *    45982    46081: gap of 100 bp
            *    46082    46844: contig of 763 bp in length
            *    46845    46944: gap of 100 bp
            *    46945    47709: contig of 765 bp in length
            *    47710    47809: gap of 100 bp
            *    47810    48550: contig of 741 bp in length
            *    48551    48650: gap of 100 bp
            *    48651    49420: contig of 770 bp in length
            *    49421    49520: gap of 100 bp
            *    49521    50284: contig of 764 bp in length
            *    50285    50384: gap of 100 bp
            *    50385    51169: contig of 785 bp in length
            *    51170    51269: gap of 100 bp
            *    51270    52054: contig of 785 bp in length
            *    52055    52154: gap of 100 bp
            *    52155    52930: contig of 776 bp in length
            *    52931    53030: gap of 100 bp
            *    53031    53803: contig of 773 bp in length
            *    53804    53903: gap of 100 bp
            *    53904    54670: contig of 767 bp in length
            *    54671    54770: gap of 100 bp
            *    54771    55546: contig of 776 bp in length
            *    55547    55646: gap of 100 bp
            *    55647    56408: contig of 762 bp in length
            *    56409    56508: gap of 100 bp
            *    56509    57273: contig of 765 bp in length
            *    57274    57373: gap of 100 bp
            *    57374    58141: contig of 768 bp in length
            *    58142    58241: gap of 100 bp
            *    58242    58983: contig of 742 bp in length
            *    58984    59083: gap of 100 bp
            *    59084    59851: contig of 768 bp in length

  Query Match             1.7%;  Score 53;  DB 14;  Length 62449;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1892 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 1944


RESULT 18
AC026771/c
LOCUS       AC026771               69808 bp    DNA     linear   PRI 13-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone CTB-134F13, complete sequence.
ACCESSION   AC026771
VERSION     AC026771.6  GI:28933549
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 69808)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 69808)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 69808)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 69808)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 69808)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 13, 2003 this sequence version replaced gi:27436786.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .69808
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTB-134F13"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 69808;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32798 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 32746


RESULT 19
AL713971
LOCUS       AL713971               70621 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-269C15 on chromosome 6 contains
            the FABGL gene for hydroxysteroid (17-beta) dehydrogenase 8, the
            RING1 gene for ring finger protein 1, a zinger finger protein
            pseudogene, a HIV TAT specific factor 1 (HTATSF1) pseudogene, a
            novel transcript, the SACML2 gene for SAC2 suppressor of actin
            mutations 2-like (yeast), the 5' end of the RPS18 gene for
            ribosomal protein S18 and four CpG islands, complete sequence.
ACCESSION   AL713971
VERSION     AL713971.7  GI:20428955
KEYWORDS    HTG; ARE-1; ARE1; CpG island; D6S2245E; DKFZp547I194; FABG; FABGL;
            H2-KE6; HIV TAT specific factor 1; HKE3; HKE6; HSD17B8; HTATSF1;
            hydroxysteroid dehydrogenase 8; KE-3; KE3; KE6; ribosomal protein
            S18; ring finger protein 1; RING1; RING2; RNF1; RP5-1033B10; RPS18;
            SAC2; SAC2 suppressor of actin mutations 2-like; SACML2; TAT-SF1;
            zinc finger protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 70621)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 3, 2002 this sequence version replaced gi:20338849.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            XXbac-269C15 is from a CHORI-502 human bac - COX cell line library
            VECTOR: pTARBAC2.1
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .70621
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-269C15"
                     /clone_lib="CHORI-502"
     misc_feature    2000
                     /note="Clone_right_end: XXbac-22I16"
     gene            join(AL662824.9:186261. .186339,AL662824.9:186521. .186738,
                     AL662824.9:186840. .186956,AL662824.9:187066. .187158,
                     AL662824.9:187259. .187344,AL662824.9:187447. .187531,
                     AL662824.9:187749. .187791,2026. .2100,2289. .2502)
                     /gene="HSD17B8"
                     /locus_tag="XXbac-BCX269C15.6-001"
     mRNA            join(AL662824.9:186261. .186339,AL662824.9:186521. .186738,
                     AL662824.9:186840. .186956,AL662824.9:187066. .187158,
                     AL662824.9:187259. .187344,AL662824.9:187447. .187531,
                     AL662824.9:187749. .187791,2026. .2100,2289. .2502)
                     /gene="HSD17B8"
                     /locus_tag="XXbac-BCX269C15.6-001"
                     /product="hydroxysteroid (17-beta) dehydrogenase 8"
                     /note="match: ESTs: AL555459.1 BF523076.1 BI159884.1
                     BM460286.1
                     match: cDNAs: AF272012.1 BC008185.1 D82061.1"
     gene            join(AL662824.9:186409. .186738,AL662824.9:186840. .186956,
                     AL662824.9:187066. .187158,AL662824.9:187259. .187344,
                     AL662824.9:187447. .187531,AL662824.9:187749. .187791,
                     2026. .2100,2289. .2401)
                     /gene="HSD17B8"
                     /locus_tag="XXbac-BCX269C15.6-002"
     mRNA            join(AL662824.9:186409. .186738,AL662824.9:186840. .186956,
                     AL662824.9:187066. .187158,AL662824.9:187259. .187344,
                     AL662824.9:187447. .187531,AL662824.9:187749. .187791,
                     2026. .2100,2289. .2401)
                     /gene="HSD17B8"
                     /locus_tag="XXbac-BCX269C15.6-002"
                     /product="hydroxysteroid (17-beta) dehydrogenase 8"
                     /note="match: ESTs: AI186776.1 AI359146.1 AI992249.1
                     AW134912.1 BQ218704.1"
     CDS             join(AL662824.9:186288. .186339,AL662824.9:186521. .186738,
                     AL662824.9:186840. .186956,AL662824.9:187066. .187158,
                     AL662824.9:187259. .187344,AL662824.9:187447. .187531,
                     AL662824.9:187749. .187791,2026. .2100,2289. .2305)
                     /gene="HSD17B8"
                     /locus_tag="XXbac-BCX269C15.6-002"
                     /standard_name="OTTHUMP00000029564"
                     /note="match: proteins: AAH08185 P50171 Q8WMN4 Q92506
                     Q96KK9"
                     /codon_start=1
                     /product="hydroxysteroid (17-beta) dehydrogenase 8"
                     /protein_id="CAI17657.1"
                     /db_xref="GI:55961772"
                     /db_xref="InterPro:IPR002198"
                     /db_xref="InterPro:IPR002347"
                     /translation="MASQLQNRLRSALALVTGAGSGIGRAVSVRLAGEGATVAACDLD
                     RAAAQETVRLLGGPGSKEGPPRGNHAAFQADVSEARAARCLLEQVQACFSRPPSVVVS
                     CAGITQDEFLLHMSEDDWDKVIAVNLKGTFLVTQAAAQALVSNGCRGSIINISSIVGK
                     VGNVGQTNYAASKAGVIGLTQTAARELGRHGIRCNSVLPGFIATPMTQKVPQKVVDKI
                     TEMIPMGHLGDPEDVADVVAFLASEDSGYITGTSVEVTGGLFM"
     gene            join(4151. .4295,4416. .4551,5277. .5437,5566. .5781,
                     6809. .7198,7380. .7653,7949. .8373)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-002"
     mRNA            join(4151. .4295,4416. .4551,5277. .5437,5566. .5781,
                     6809. .7198,7380. .7653,7949. .8373)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-002"
                     /product="ring finger protein 1"
                     /note="match: ESTs: AI208750.1 AL535164.1 BI259576.1
                     match: cDNAs: BC002922.1 BC009070.1"
     CDS             join(4474. .4551,5277. .5437,5566. .5781,6809. .7198,
                     7380. .7653,7949. .8050)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-002"
                     /standard_name="OTTHUMP00000029850"
                     /note="match: proteins: AAH02922 Q8WMN5 Q921Z8 Q96KL0
                     Q9VB08"
                     /codon_start=1
                     /product="ring finger protein 1"
                     /protein_id="CAI17650.1"
                     /db_xref="GI:55961767"
                     /db_xref="GOA:Q86V19"
                     /db_xref="InterPro:IPR001841"
                     /db_xref="UniProt/TrEMBL:Q86V19"
                     /translation="MTTPANAQNASKTWELSLYELHRTPQEAIMDGTEIAVSPRSLHS
                     ELMCPICLDMLKNTMTTKECLHRFCSDCIVTALRSGNKECPTCRKKLVSKRSLRPDPN
                     FDALISKIYPSREEYEAHQDRVLIRLSRLHNQQALSSSIEEGLRMQAMHRAQRVRRPI
                     PGSDQTTTMSGGEGEPGEGEGDGEDVSSDSAPDSAPGPAPKRPRGGGAGGSSVGTGGG
                     GTGGVGGGAGSEDSGDRGGTLGGGTLGPPSPPGAPSPPEPGGEIELVFRPHPLLVEKG
                     EYCQTRYVKTTGNATVDHLSKYLALRIALERRQQQEAGEPGGPGGGASDTGGPDGCGG
                     EGGGAGGGDGPEEPALPSLEGVSEKQYTIYIAPGGGAFTTLNGSLTLELVNEKFWKVS
                     RPLELCYAPTKDPK"
     gene            join(5211. .5437,5566. .5781,6809. .7198,7380. .7653,
                     7949. .8373)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-001"
     mRNA            join(5211. .5437,5566. .5781,6809. .7198,7380. .7653,
                     7949. .8373)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-001"
                     /product="ring finger protein 1"
                     /note="match: cDNAs: Z14000.1"
     CDS             join(5286. .5437,5566. .5781,6809. .7198,7380. .7653,
                     7949. .8050)
                     /gene="RING1"
                     /locus_tag="XXbac-BCX269C15.3-001"
                     /standard_name="OTTHUMP00000029849"
                     /note="match: proteins: O35730 Q921Z8"
                     /codon_start=1
                     /product="ring finger protein 1"
```

```
                     /protein_id="CAI17651.1"
                     /db_xref="GI:55961768"
                     /db_xref="InterPro:IPR001841"
                     /translation="MDGTEIAVSPRSLHSELMCPICLDMLKNTMTTKECLHRFCSDCI
                     VTALRSGNKECPTCRKKLVSKRSLRPDPNFDALISKIYPSREEYEAHQDRVLIRLSRL
                     HNQQALSSSIEEGLRMQAMHRAQRVRRPIPGSDQTTTMSGGEGEPGEGEGDGEDVSSD
                     SAPDSAPGPAPKRPRGGGAGGSSVGTGGGGTGGVGGGAGSEDSGDRGGTLGGGTLGPP
                     SPPGAPSPPEPGGEIELVFRPHPLLVEKGEYCQTRYVKTTGNATVDHLSKYLALRIAL
                     ERRQQQEAGEPGGPGGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQYTI
                     YIAPGGGAFTTLNGSLTLELVNEKFWKVSRPLELCYAPTKDPK"
     gene            11366. .11973
                     /locus_tag="XXbac-BCX269C15.2-001"
                     /pseudo
     CDS             11366. .11973
                     /locus_tag="XXbac-BCX269C15.2-001"
                     /note="match: proteins: P28698 Q96H95 Q96I71 Q9NRY0"
                     /pseudo
                     /codon_start=1
                     /product="zinc finger protein 314 pseudogene"
     gene            join(33490. .34019,34322. .34431,35078. .35432)
                     /locus_tag="XXbac-BCX269C15.1-003"
                     /pseudo
     CDS             join(33490. .34019,34322. .34431,35078. .35432)
                     /locus_tag="XXbac-BCX269C15.1-003"
                     /note="match: proteins: O43719 Q99730 Q9CT41 Q9DAU3"
                     /pseudo
                     /codon_start=1
                     /product="HIV TAT specific factor 1 pseudogene"
     gene            join(45181. .45302,45437. .45536,46263. .46358,47160. .47353,
                     50345. .50531)
                     /locus_tag="XXbac-BCX269C15.5-001"
     gene            join(45181. .45302,45437. .45536,46263. .46358,47225. .47353,
                     50345. .50531)
                     /locus_tag="XXbac-BCX269C15.5-002"
     mRNA            join(45181. .45302,45437. .45536,46263. .46358,47160. .47353,
                     50345. .50531)

  Query Match             1.7%;  Score 53;  DB 8;  Length 70621;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34070 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 34122


RESULT 20
AC015910
LOCUS       AC015910               90476 bp    DNA     linear   PRI 01-APR-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-690G19, complete sequence.
ACCESSION   AC015910
VERSION     AC015910.10  GI:19852158
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 90476)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-690G19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 90476)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 90476)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 90476)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 1, 2002 this sequence version replaced gi:19683792.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L832
                Center clone name: 690_G_19
            --------------
```

```
          Only the first 90.5 kilobases of this clone are being submitted.
          The remainder of the clone is overlapped by either accession number
          AC005288 [WICGR project L301] or accession
          number AC004408 [WICGR project L309].
FEATURES             Location/Qualifiers
     source          1. .90476
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-690G19"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   111. .405
                     /rpt_family="AluJb"
     repeat_region   complement(572. .750)
                     /rpt_family="AluSg/x"
     repeat_region   complement(752. .920)
                     /rpt_family="AluSx"
     repeat_region   complement(921. .1140)
                     /rpt_family="AluSg/x"
     repeat_region   complement(1141. .1267)
                     /rpt_family="AluSx"
     repeat_region   complement(1278. .1559)
                     /rpt_family="AluJb"
     repeat_region   complement(1660. .1947)
                     /rpt_family="AluJo"
     repeat_region   complement(2011. .2302)
                     /rpt_family="AluSx"
     repeat_region   2330. .2625
                     /rpt_family="AluSx"
     unsure          2378. .2386
                     /note="extensive edit here"
     repeat_region   2749. .3024
                     /rpt_family="AluSx"
     repeat_region   3049. .3126
                     /rpt_family="(CATATA)n"
     repeat_region   complement(3393. .3699)
                     /rpt_family="AluSp"
     repeat_region   complement(4229. .4354)
                     /rpt_family="FLAM_A"
     repeat_region   4471. .4521
                     /rpt_family="Alu"
     repeat_region   complement(4522. .4785)
                     /rpt_family="AluSg"
     repeat_region   4793. .4919
                     /rpt_family="(TTCC)n"
     repeat_region   complement(4931. .5209)
                     /rpt_family="AluSx"
     repeat_region   5465. .5772
                     /rpt_family="AluSx"
     repeat_region   complement(6691. .6796)
                     /rpt_family="FLAM_C"
     repeat_region   complement(6833. .6928)
                     /rpt_family="AluJo"
     repeat_region   complement(6963. .6999)
                     /rpt_family="MIR"
     repeat_region   7000. .7287
                     /rpt_family="AluJo"
     repeat_region   complement(7288. .7338)
                     /rpt_family="MIR"
     repeat_region   complement(7339. .7628)
                     /rpt_family="AluSg"
     repeat_region   complement(7629. .7752)
                     /rpt_family="MIR"
     repeat_region   7859. .8138
                     /rpt_family="AluSx"
     repeat_region   complement(8415. .8748)
                     /rpt_family="AluSx"
     repeat_region   complement(8772. .9067)
                     /rpt_family="AluSc"
     repeat_region   complement(9082. .9207)
                     /rpt_family="FLAM_C"
     repeat_region   complement(9550. .9858)
                     /rpt_family="AluSx"
     repeat_region   complement(10613. .10908)
                     /rpt_family="AluSx"
     repeat_region   11112. .11147
                     /rpt_family="AT_rich"
     repeat_region   11261. .11376
                     /rpt_family="AluJo"
     repeat_region   11377. .11681
                     /rpt_family="AluSp"
     repeat_region   11682. .11996
                     /rpt_family="AluY"
     repeat_region   11997. .12164
                     /rpt_family="AluJo"
     repeat_region   complement(12789. .13093)
                     /rpt_family="AluSq"
     repeat_region   complement(13819. .13888)
                     /rpt_family="AluJo/FRAM"
     repeat_region   complement(13889. .14203)
                     /rpt_family="AluSx"
     repeat_region   complement(14204. .14303)
                     /rpt_family="AluJo/FRAM"
     repeat_region   14528. .14656
                     /rpt_family="AluSx"
     repeat_region   14657. .14957
                     /rpt_family="AluSx"
     repeat_region   15137. .15372
                     /rpt_family="AluJo"
     repeat_region   15389. .15657

  Query Match             1.7%;  Score 53;  DB 8;  Length 90476;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14575 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 14627


RESULT 21
AP000208
LOCUS       AP000208              100000 bp    DNA     linear   PRI 08-JAN-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
            clone f43D11-119B8, segment 6/12, complete sequence.
ACCESSION   AP000208
VERSION     AP000208.1  GI:4827146
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Homo sapiens 1,109,292bp genomic DNA of 21q22.1 (REGION:
            D21S226-AML CLONE RANGE: f43D11-119B8)
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 100000)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     E. coli transposon insertion:The present data does not contain E.
            coli transposon sequences which integrated in the
            original/previous sequences.  We determined the boundary between
            the insertion and genomic sequences experimentally, removed the
            insertion sequences, reconstituted the present data. The sequencing
```

```
            project is supported by Japan Science Technology Corporation (JST)
            and The Institute of Physical and Chemical Research (RIKEN).
FEATURES             Location/Qualifiers
     source          1. .100000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 100000;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33384 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 33436


RESULT 22
AL844527
LOCUS       AL844527              109523 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone DAQB-314F24 on chromosome 6 Contains
            the 5' end of the COL11A2 gene for collagen type XI alpha 2, the
            RXRB gene for retinoid X receptor, beta, the SLC39A7 gene for
            solute carrier family 39 (zinc transporter), member 7, the HSD17B8
            gene for hydroxysteroid (17-beta) dehydrogenase 8, the RING1 gene
            for ring finger protein 1, the ZNF314P pseudogene for zinc finger
            protein 314, the HTATSF1P pseudogene for HIV TAT specific factor 1,
            the HCG25 gene for HLA complex group 25, the VPS52 gene for
            vacuolar protein sorting 52 (yeast), the RPS18 gene for ribosomal
            protein S18, the B3GALT4 gene for UDP-Gal:betaGlcNAc beta
            1,3-galactosyltransferase, polypeptide 4, the 3'end of the C6orf11
            gene for chromosome 6 open reading frame 11, and 7 CpG islands,
            complete sequence.
ACCESSION   AL844527
VERSION     AL844527.6  GI:24474476
KEYWORDS    HTG; B3GALT4; C6ORF11; COL11A2; CpG island; HCG25; HSD17B8;
            HTATSF1P; RING1; RPS18; RXRB; SLC39A7; VPS52; ZNF314P.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 109523)
  AUTHORS   Leongamornlert,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 1, 2002 this sequence version replaced gi:22204803.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            DAQB-314F24 is from a DNA-arts QBL human bac library VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .109523
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAQB-314F24"
                     /clone_lib="DNA-arts-BAC.1-QBL.1"
     misc_feature    1
                     /note="Clone_left_end: DAQB-314F24"
     gene            join(complement(2001. .20909),
                     complement(AL845446.7:82994. .93842))
                     /gene="COL11A2"
                     /locus_tag="DAQB-79P13.8-001"
     mRNA            join(complement(20615. .20909),complement(17776. .17925),
                     complement(17434. .17644),complement(16818. .16980),
                     complement(15083. .15274),complement(13448. .13510),
                     complement(9568. .9627),complement(9421. .9462),
                     complement(9125. .9187),complement(8715. .8789),
                     complement(8176. .8262),complement(7888. .7944),
                     complement(7688. .7741),complement(7530. .7583),
                     complement(7377. .7430),complement(7136. .7189),
                     complement(6892. .6945),complement(6765. .6809),
                     complement(6589. .6642),complement(6119. .6163),
                     complement(5869. .5922),complement(5638. .5682),
                     complement(5449. .5502),complement(5178. .5222),
                     complement(4890. .4943),complement(4716. .4760),
                     complement(4473. .4526),complement(4031. .4138),
                     complement(2984. .3037),complement(2622. .2675),
                     complement(2468. .2512),complement(2330. .2383),
                     complement(2158. .2202),complement(2001. .2012),
                     complement(AL845446.7:93801. .93842),
                     complement(AL845446.7:93647. .93700),
                     complement(AL845446.7:93359. .93412),
                     complement(AL845446.7:92829. .92936),
                     complement(AL845446.7:92588. .92677),
                     complement(AL845446.7:92349. .92402),
                     complement(AL845446.7:92012. .92119),
                     complement(AL845446.7:91766. .91873),
                     complement(AL845446.7:91570. .91623),
                     complement(AL845446.7:91413. .91466),
                     complement(AL845446.7:91110. .91217),
                     complement(AL845446.7:90857. .90910),
                     complement(AL845446.7:90626. .90679),
                     complement(AL845446.7:90347. .90400),
                     complement(AL845446.7:90131. .90184),
                     complement(AL845446.7:89683. .89790),
                     complement(AL845446.7:89255. .89308),
                     complement(AL845446.7:89005. .89058),
                     complement(AL845446.7:88818. .88871),
                     complement(AL845446.7:88098. .88151),
                     complement(AL845446.7:87725. .87832),
                     complement(AL845446.7:87563. .87616),
                     complement(AL845446.7:87369. .87422),
                     complement(AL845446.7:87019. .87126),
                     complement(AL845446.7:86812. .86865),
                     complement(AL845446.7:86470. .86505),
                     complement(AL845446.7:86230. .86283),
                     complement(AL845446.7:85848. .86115),
                     complement(AL845446.7:85151. .85263),
                     complement(AL845446.7:84566. .84772),
                     complement(AL845446.7:82994. .84117))
                     /gene="COL11A2"
                     /locus_tag="DAQB-79P13.8-001"
                     /product="collagen type XI alpha 2"
                     /note="match: ESTs: AU118676.1 BU164080.1 BX282345.1
                     match: cDNAs: J04974.1 L18987.1 U16789.1"
     CDS             join(complement(20615. .20696),complement(17776. .17925),
                     complement(17434. .17644),complement(16818. .16980),
                     complement(15083. .15274),complement(13448. .13510),
```

```
complement(9568. .9627),complement(9421. .9462),
complement(9125. .9187),complement(8715. .8789),
complement(8176. .8262),complement(7888. .7944),
complement(7688. .7741),complement(7530. .7583),
complement(7377. .7430),complement(7136. .7189),
complement(6892. .6945),complement(6765. .6809),
complement(6589. .6642),complement(6119. .6163),
complement(5869. .5922),complement(5638. .5682),
complement(5449. .5502),complement(5178. .5222),
complement(4890. .4943),complement(4716. .4760),
complement(4473. .4526),complement(4031. .4138),
complement(2984. .3037),complement(2622. .2675),
complement(2468. .2512),complement(2330. .2383),
complement(2158. .2202),complement(2001. .2012),
complement(AL845446.7:93801. .93842),
complement(AL845446.7:93647. .93700),
complement(AL845446.7:93359. .93412),
complement(AL845446.7:92829. .92936),
complement(AL845446.7:92588. .92677),
complement(AL845446.7:92349. .92402),
complement(AL845446.7:92012. .92119),
complement(AL845446.7:91766. .91873),
complement(AL845446.7:91570. .91623),
complement(AL845446.7:91413. .91466),
complement(AL845446.7:91110. .91217),
complement(AL845446.7:90857. .90910),
complement(AL845446.7:90626. .90679),
complement(AL845446.7:90347. .90400),
complement(AL845446.7:90131. .90184),
complement(AL845446.7:89683. .89790),
complement(AL845446.7:89255. .89308),
complement(AL845446.7:89005. .89058),
complement(AL845446.7:88818. .88871),
complement(AL845446.7:88098. .88151),
complement(AL845446.7:87725. .87832),
complement(AL845446.7:87563. .87616),
complement(AL845446.7:87369. .87422),
complement(AL845446.7:87019. .87126),
complement(AL845446.7:86812. .86865),
complement(AL845446.7:86470. .86505),
complement(AL845446.7:86230. .86283),
complement(AL845446.7:85848. .86115),
complement(AL845446.7:85151. .85263),
complement(AL845446.7:84566. .84772),
complement(AL845446.7:83977. .84117))
/gene="COL11A2"
/locus_tag="DAQB-79P13.8-001"
/standard_name="OTTHUMP00000014556"
/note="match: proteins: Q28083 Q8IZC6 Q8UUM5 Q9JLI2"
/codon_start=1
/product="collagen type XI alpha 2"
/protein_id="CAI41834.1"
/db_xref="GI:57209737"
/db_xref="GOA:Q5JP94"
/db_xref="InterPro:IPR000885"
/db_xref="InterPro:IPR001791"
/db_xref="InterPro:IPR003129"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/TrEMBL:Q5JP94"
/translation="MERCSRCHRLLLLLPLVLGLSAAPGWAGAPPVDVLRALRFPSLP
DGVRRAKGICPADVAYRVARPAQLSAPTRQLFPGGFPKDFSLLTVVRTRPGLQAPLLT
LYSAQGVRQLGLELGRPVRFLYEDQTGRPQPPSQPVFRGLSLADGKWHRVAVAVKGQS
VTLIVDCKKRVTRPLPRSARPVLDTHGVIIFGARILDEEVFEGDVQELAIVPGVQAAY
ESCEQKELECEGGQRERPQNQQPHRAQRSPQQQPSRLHRPQNQEPQSQDPTPGEEEEI
LESSLLPPLEEAAHGPRGLKGEKGEPAVLEPGMLVEGPPGPEGPAGLIGPPGIQGNPG
PVGDPGERGPPGRAGLPGSDGAPGPPGTSLMLPFRFGSGGGDKGPVVAAQEAQAQAIL
QQARLALRGPPGPMGYTGRPGPLGQPGSPGLKGESGDLGPQGPRGPQGLTGPPGKAGR
RGRAGADGARGMPGDPGVKGDRGFDGLPGLPGEKGHRGDTGAQGLPGPPGEDGERGDD
GEIGPRGLPGESGPRGLLGPKGPPGIPGPPGVRGMDGPQGPKGSLGPQGEPGPPGQQG
TPGTQGLPGPQGAIGPHGEKGPQGKPGLPGMPGSDGPPGHPGKEGPPGTKGNQGPSGP
QGPLGYPGPRGVKGVDGIRGLKGHKGEKGEDGFPGFKGDIGVKGDRGEVGVPGSRGED
GPEGPKGRTGPTGDPGPPGLMGEKGKLGVPGLPGYPGRQGPKGSLGFPGFPGASGEKG
```

```
ARGLSGKSGPRGERGPTGPRGQRGPRGATGKSGAKGTSGGDGPHGPPGERGLPGPQGP
NGFPGPKGPPGPPGKDGLPGHPGQRGEVGFQGKTGPPGPPGVVGPQGAAGETGPMGER
GHPGPPGPPGEQGLPGTAGKEGTKGDPGPPGAPGKDGPAGLRGPPGERGLPGTAGGPG

  Query Match             1.7%;  Score 53;  DB 8;  Length 109523;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66876 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 66928


RESULT 23
AL360268/c
LOCUS       AL360268              110130 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-379C10 on chromosome 9 Contains
            the gene for a novel protein, the 5' end of the PCSCL gene for
            likely ortholog of rat peroxisomal Ca-dependent solute carrier-like
            protein (KIAA1896), two novel genes and a CpG island, complete
            sequence.
ACCESSION   AL360268
VERSION     AL360268.14  GI:19351891
KEYWORDS    HTG; CpG island; KIAA1896; PCSCL.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110130)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 11, 2002 this sequence version replaced gi:18650682.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-379C10 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .110130
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIB753C10379"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-379C10"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-379C10"
```

```
misc_feature    303
                /note="Clone_right_end: RP11-203J24"
gene            complement(join(48570. .48742,66591. .66797))
                /locus_tag="RP11-379C10.1-001"
mRNA            complement(join(48570. .48742,66591. .66797))
                /locus_tag="RP11-379C10.1-001"
                /note="match: ESTs: AA961281.1"
misc_feature    64768
                /note="Tandem repeat. Forced join. Gap size estimated to
                be approximately 200bp by restriction digest data."
gene            complement(join(70668. .73335,76026. .76756))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-001"
mRNA            complement(join(70668. .73335,76026. .76756))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-001"
                /note="match: ESTs: AA043145.1 AA340072.1 AA359485.1
                AA381105.1 AA631602.1 AA632403.1 AI081135.1 AI147614.1
                AI417373.1 AI831258.1 AI864878.1 AW082728.1 AW449992.1
                AW769468.1 AW838390.1 AW956385.1 AW963894.1 AW963898.1
                BE044419.1 BE293321.1 BE385388.1 BE741976.1 BF131845.1
                BF724763.1 BF724764.1 BF796375.1 BG009209.1 BG340504.1
                BM016964.1 BM464251.1 BM510565.1 BM511394.1 BM543547.1
                BM678197.1 BM725183.1 BM740427.1 BM740504.1 BM740522.1
                BM750316.1 BM757831.1 BM759701.1 BM853143.1 BM905218.1
                BM914393.1 BM931058.1 BQ231034.1 BQ441346.1 BU059062.1
                BU162007.1 CA310459.1 CA312400.1 W69444.1
                match: cDNAs: AB060227.1 AL833392.1"
gene            complement(join(70668. .71261,72809. .73177))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
mRNA            complement(join(70668. .71261,72809. .73177))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
                /note="match: ESTs: AA281868.1 BG008293.1 BG912959.1
                BM684531.1 BQ370637.1"
polyA_site      complement(70668)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
polyA_site      complement(70672)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
polyA_site      complement(70674)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
polyA_signal    complement(70691. .70696)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
gene            complement(join(71425. .73335,75582. .75719))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-002"
mRNA            complement(join(71425. .73335,75582. .75719))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-002"
                /note="match: ESTs: BE747080.1
                match: cDNAs: BC013371.1"
polyA_site      complement(72723)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
polyA_site      complement(72731)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
polyA_signal    complement(72743. .72748)
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-004"
CDS             complement(join(72863. .73335,76026. .76536))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-001"
                /standard_name="OTTHUMP00000022227"
                /note="match: proteins: Q8WU12 Q9BE21"
                /codon_start=1
                /protein_id="CAH73132.1"
                /db_xref="GI:55664183"
                /db_xref="UniProt/TrEMBL:Q69YI7"
                /translation="MAVPAKKRKMNFSEREVEIIVEELELKKHLLVNHFNAGVPLAAK
                SAAWHGILRRVNAVATCRRELPEVKKKWSDLKTEVRRKVAQVRAAVEGGEAPGPTEED
                GAGGPGTGGGSGGGGPAVAPVLLTPMQQRICNLLGEATIISLPSTTEIHPVALGPSAT
                AAAATVTLTQIPTETTYHTLEEGVVEYCTAEAPPPLPPETPVDMMAQHADTSVKPQAL
                KSRIALNSAKLIQEQRVTNLHVKEIAQHLEQQNDLLQMIRRSQEVQACAQERQAQAME
                GTQAALSVLIQVLRPMIKDFRRYLQSNTANPAPASDPGQVAQNGQPDSIIQ"
gene            complement(join(73077. .73335,77558. .77641))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-005"
mRNA            complement(join(73077. .73335,77558. .77641))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-005"
                /note="match: ESTs: AA393124.1"
gene            complement(join(73520. .74191,75582. .76601))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-003"
mRNA            complement(join(73520. .74191,75582. .76601))
                /gene="RP11-379C10.2"
                /locus_tag="RP11-379C10.2-003"
                /note="match: ESTs: BG774313.1
                match: cDNAs: BC021580.1"
gene            join(77636. .78015,AL590708.18:2401. .2527,
                AL590708.18:2614. .2701,AL590708.18:3675. .3786,
                AL590708.18:4986. .5144,AL590708.18:7034. .7186,
                AL590708.18:7455. .>7494)
                /gene="RP11-395P17.4"
                /locus_tag="RP11-395P17.4-004"
mRNA            join(77636. .78015,AL590708.18:2401. .2527,
                AL590708.18:2614. .2701,AL590708.18:3675. .3786,
                AL590708.18:4986. .5144,AL590708.18:7034. .7186,
                AL590708.18:7455. .>7494)
                /gene="RP11-395P17.4"
                /locus_tag="RP11-395P17.4-004"
                /note="match: ESTs: AL523261.1 AL555731.1 AL559911.1
                BF335109.1 BF376153.1 BF376154.1 BF799713.1"
gene            join(77728. .78015,AL590708.18:2401. .2527,
                AL590708.18:2614. .2701,AL590708.18:3385. .3420,
                AL590708.18:3675. .3786,AL590708.18:4986. .5144,
                AL590708.18:7034. .7186,AL590708.18:7455. .7562,
                AL590708.18:7658. .7825,AL590708.18:8334. .8484,
                AL590708.18:8565. .10550)
                /gene="RP11-395P17.4"
                /locus_tag="RP11-395P17.4-003"
mRNA            join(77728. .78015,AL590708.18:2401. .2527,
                AL590708.18:2614. .2701,AL590708.18:3385. .3420,
                AL590708.18:3675. .3786,AL590708.18:4986. .5144,
                AL590708.18:7034. .7186,AL590708.18:7455. .7562,
                AL590708.18:7658. .7825,AL590708.18:8334. .8484,
                AL590708.18:8565. .10550)
                /gene="RP11-395P17.4"
                /locus_tag="RP11-395P17.4-003"
                /note="match: cDNAs: BC037109.1"
CDS             join(77755. .78015,AL590708.18:2401. .2527,
                AL590708.18:2614. .2701,AL590708.18:3675. .3786,
                AL590708.18:4986. .5144,AL590708.18:7034. .7186,
                AL590708.18:7455. .>7494)
                /gene="RP11-395P17.4"
                /locus_tag="RP11-395P17.4-004"
                /standard_name="OTTHUMP00000022252"
                /codon_start=1
                /protein_id="CAH73133.1"
                /db_xref="GI:55664184"
                /db_xref="GOA:Q5VWU4"
                /db_xref="InterPro:IPR001993"
                /db_xref="InterPro:IPR002048"
                /db_xref="InterPro:IPR002067"
                /db_xref="InterPro:IPR002167"
                /db_xref="UniProt/TrEMBL:Q5VWU4"

Query Match             1.7%;  Score 53;  DB 8;  Length 110130;
Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29751 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 29699


RESULT 24
AP000247
LOCUS       AP000247              113919 bp    DNA     linear   PRI 15-MAR-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 21q, clone:RP5-871G18,
            complete sequence.
ACCESSION   AP000247
VERSION     AP000247.2  GI:14861122
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (1999)
REFERENCE   2  (bases 1 to 113919)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jul 17, 2001 this sequence version replaced gi:4835616.
FEATURES             Location/Qualifiers
     source          1. .113919
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="RP5-871G18"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 113919;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13942 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 13994


RESULT 25
HSJ733M16
LOCUS       HSJ733M16             120825 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP4-733M16 on chromosome
            1p36.11-36.23 Contains four novel genes, a novel gene (FLJ33962), a
            novel gene (MGC10731), the 3' end of a novel gene (KIAA1332), a
            metallothionein family pseudogene and three CpG islands, complete
            sequence.
ACCESSION   AL109627
VERSION     AL109627.18  GI:6742145
KEYWORDS    HTG; FLJ33962; KIAA1332; MGC10731.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 120825)
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 24, 2000 this sequence version replaced gi:6723665.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP4-733M16 is from the library RPCI-4 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .120825
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M16733"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p36.11-36.23"
                     /clone="RP4-733M16"
                     /clone_lib="RPCI-4"
     misc_feature    1
                     /note="Clone_left_end: RP4-733M16"
     gene            join(complement(AL358794.19:480. .678),10770. .10886,
                     21772. .24375)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-002"
     mRNA            join(complement(AL358794.19:480. .678),10770. .10886,
                     21772. .24375)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-002"
                     /product="F-box protein 42"
                     /note="match: cDNAs: Em:AL833207.1"
     misc_feature    33887
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33888
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33892
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33903
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33910
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33921
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33933
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33938
```

```
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    33942
                     /note="Possible local misassembly of alu containing reads.
                     Overall assembly confirmed by restriction digest."
     misc_feature    44370. .44374
                     /note="Single clone region. Within alu region. pUC read
                     pairs confirm assembly."
     gene            join(63096. .63133,63508. .63593,64285. .64401,64904. .>65607)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-004"
     mRNA            join(63096. .63133,63508. .63593,64285. .64401,64904. .>65607)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-004"
                     /product="F-box protein 42"
                     /note="match: ESTs: Em:Z41965.1"
     gene            join(63262. .63593,64285. .64401,64904. .>65607)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-003"
     mRNA            join(63262. .63593,64285. .64401,64904. .>65607)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-003"
                     /product="F-box protein 42"
                     /note="match: ESTs: Em:AL534909.2"
     CDS             join(63519. .63593,64285. .64401,64904. .>65607)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-003"
                     /standard_name="OTTHUMP00000002237"
                     /codon_start=1
                     /product="F-box protein 42"
                     /protein_id="CAI22561.1"
                     /db_xref="GI:56203757"
                     /db_xref="GOA:Q6P3S6"
                     /db_xref="UniProt/Swiss-Prot:Q6P3S6"
                     /translation="MSSPPQIVIDDATILILGGCGGPNALFKDAWLLHMHSGPWAWQP
                     LKVENEEHGAPELWCHPACRVGQCVVVFSQAPSGRAPLSPSLNSRPSPISATPPALVP
                     ETREYRSQSPVRSMDEAPCVNGRWGTLRPRAQRQTPSGSREGSLSPARGDGSPILNGG
                     SLSPGTAAVGGSSLDSPVQAISPSTPSAPEGYDLKIGLSLAPRRGSLPDQKDLRLGSI
                     DLNWDLKPASSSNPMDGMDNRTVGGSMRHPPEQTNGVHTPPHVASALAGAVSPGALRR
                     SLEAIKAMSSKGPSASAALSPPL"
     gene            join(complement(AL358794.19:37566. .37765),
                     complement(AL358794.19:480. .746),10770. .10886,
                     21772. .21906,59930. .60083,60868. .60978,62958. .63054,
                     63537. .63593,64285. .64401,64904. .69850)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-001"
     mRNA            join(complement(AL358794.19:37566. .37765),
                     complement(AL358794.19:480. .746),10770. .10886,
                     21772. .21906,59930. .60083,60868. .60978,62958. .63054,
                     63537. .63593,64285. .64401,64904. .69850)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-001"
                     /product="F-box protein 42"
                     /note="match: cDNAs: Em:AB037753.1 Em:AK055598.1"
     CDS             join(complement(AL358794.19:480. .729),10770. .10886,
                     21772. .21906,59930. .60083,60868. .60978,62958. .63054,
                     63537. .63593,64285. .64401,64904. .66019)
                     /gene="FBXO42"
                     /locus_tag="RP4-733M16.6-001"
                     /standard_name="OTTHUMP00000002236"
                     /note="match: proteins: Tr:Q86XI0 Tr:Q8CE76"
                     /codon_start=1
                     /product="F-box protein 42"
                     /protein_id="CAI22560.1"
                     /db_xref="GI:56203756"
                     /db_xref="GOA:Q6P3S6"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="InterPro:IPR006652"
                     /db_xref="InterPro:IPR011043"
                     /db_xref="InterPro:IPR011498"
                     /db_xref="UniProt/Swiss-Prot:Q6P3S6"
                     /translation="MASSSDSEDDSFMAVDQEETVLEGTMDQDEEPHPVLEAEETRHN
                     RSMSELPEEVLEYILSFLSPYQEHKTAALVCKQWYRLIKGVAHQCYHGFMKAVQEGNI
                     QWESRTYPYPGTPITQRFSHSACYYDANQSMYVFGGCTQSSCNAAFNDLWRLDLNSKE
                     WIRPLASGSYPSPKAGATLVVYKDLLVLFGGWTRPSPYPLHQPERFFDEIHTYSPSKN
                     WWNCIVTTHGPPPMAGHSSCVIDDKMIVFGGSLGSRQMSNDVWVLDLEQWAWSKPNIS
                     GPSPHPRGGQSQIVIDDATILILGGCGGPNALFKDAWLLHMHSGPWAWQPLKVENEEH
                     GAPELWCHPACRVGQCVVVFSQAPSGRAPLSPSLNSRPSPISATPPALVPETREYRSQ
                     SPVRSMDEAPCVNGRWGTLRPRAQRQTPSGSREGSLSPARGDGSPILNGGSLSPGTAA
                     VGGSSLDSPVQAISPSTPSAPEGYDLKIGLSLAPRRGSLPDQKDLRLGSIDLNWDLKP
                     ASSSNPMDGMDNRTVGGSMRHPPEQTNGVHTPPHVASALAGAVSPGALRRSLEAIKAM
                     SSKGPSASAALSPPLGSSPGSPGSQSLSSGETVPIPRPGPAQGDGHSLPPIARRLGHH
                     PPQSLNVGKPLYQSMNCKPMQMYVLDIKDTKEKGRVKWKVFNSSSVVGPPETSLHTVV
                     QGRGELIIFGGLMDKKQNVKYYPKTNALYFVRAKR"
     gene            75291. .75476
                     /locus_tag="RP4-733M16.7-001"
                     /pseudo
     CDS             75291. .75476
                     /locus_tag="RP4-733M16.7-001"
                     /note="match: proteins: Sw:O19000 Sw:P02795 Sw:P02799
                     Sw:P04355 Sw:P04732 Sw:P14425 Sw:P79379 Sw:P80297
                     Sw:Q9XST5 Tr:Q8WVB5"
                     /pseudo
                     /codon_start=1
                     /product="metallothionein family pseudogene"
     gene            join(79689. .79846,87643. .88014)
                     /gene="RP4-733M16.4"
                     /locus_tag="RP4-733M16.4-002"
     mRNA            join(79689. .79846,87643. .88014)
                     /gene="RP4-733M16.4"
                     /locus_tag="RP4-733M16.4-002"
                     /product="novel protein"
                     /note="match: ESTs: Em:BF001156.1 Em:CB048455.1
                     Em:CB048456.1"
     gene            join(79738. .80058,82925. .83077,83672. .83796,84043. .84179,
                     84392. .84989)
                     /gene="RP4-733M16.4"
                     /locus_tag="RP4-733M16.4-001"
     mRNA            join(79738. .80058,82925. .83077,83672. .83796,84043. .84179,

  Query Match             1.7%;  Score 53;  DB 8;  Length 120825;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52626 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 52678


RESULT 26
AC095044/c
LOCUS       AC095044               134184 bp    DNA     linear   PRI 15-OCT-2003
DEFINITION  Homo sapiens BAC clone RP11-132B16 from 7, complete sequence.
ACCESSION   AC095044 AC023954
VERSION     AC095044.3  GI:18464270
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 134184)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 134184)
  AUTHORS   Cotton,M. and VanBrunt,A.
  TITLE     The sequence of Homo sapiens BAC clone RP11-132B16
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 134184)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
```

```
REFERENCE   4  (bases 1 to 134184)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 134184)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:16418233.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0132B16
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .134184
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP11-132B16"
                     /clone_lib="RPCI-11"
     repeat_region   381. .697
                     /rpt_family="Alu"
     repeat_region   712. .1032
                     /rpt_family="Alu"
     repeat_region   1033. .1060
                     /rpt_family="AT_rich"
     repeat_region   1276. .1592
                     /rpt_family="MaLR"
     repeat_region   1594. .1772
                     /rpt_family="(TTCC)n"
     repeat_region   1775. .1852
                     /rpt_family="CT-rich"
     repeat_region   1853. .1985
                     /rpt_family="L2"
     repeat_region   2062. .2086
                     /rpt_family="Alu"
     repeat_region   2087. .2109
                     /rpt_family="(TTAA)n"
     repeat_region   2110. .2260
                     /rpt_family="Alu"
     repeat_region   2261. .2565
                     /rpt_family="Alu"
     repeat_region   2566. .2704
                     /rpt_family="Alu"
     repeat_region   2858. .2887
                     /rpt_family="(TTCA)n"
     repeat_region   3064. .3375
                     /rpt_family="ERV1"
     repeat_region   3381. .3471
                     /rpt_family="ERV1"
     repeat_region   3600. .3991
                     /rpt_family="L2"
     repeat_region   4326. .4504
                     /rpt_family="(TGAA)n"
     repeat_region   4517. .4597
                     /rpt_family="(TGAA)n"
     repeat_region   4613. .4996
                     /rpt_family="L1"
     repeat_region   5710. .5779
                     /rpt_family="MIR"
     repeat_region   6510. .6573
                     /rpt_family="(GGAA)n"
     repeat_region   6634. .6937
                     /rpt_family="Alu"
     repeat_region   7031. .7083
                     /rpt_family="A-rich"
     repeat_region   7095. .7621
                     /rpt_family="ERVL"
     repeat_region   7436. .7642
                     /rpt_family="ERV1"
     repeat_region   7653. .8019
                     /rpt_family="MaLR"
     repeat_region   8223. .8536
                     /rpt_family="Alu"
     repeat_region   8599. .8876
                     /rpt_family="Alu"
     repeat_region   8877. .8903
                     /rpt_family="(TTAA)n"
     repeat_region   9039. .9092
                     /rpt_family="(TTCC)n"
     repeat_region   9093. .9259
                     /rpt_family="(TTTC)n"
     repeat_region   9260. .9556
                     /rpt_family="Alu"
     repeat_region   10048. .10363
                     /rpt_family="Alu"
     repeat_region   10364. .10394
                     /rpt_family="(GA)n"
     repeat_region   10566. .10725
                     /rpt_family="ERV1"
     repeat_region   10729. .10908
                     /rpt_family="Alu"
     repeat_region   10909. .11207
                     /rpt_family="Alu"
     repeat_region   11209. .11470
                     /rpt_family="ERV1"
     repeat_region   11638. .12115
                     /rpt_family="MER1_type"
     repeat_region   12182. .12491
                     /rpt_family="Alu"
     repeat_region   12519. .12696
                     /rpt_family="MIR"
     repeat_region   13006. .13037
```

```
                     /rpt_family="AT_rich"
     repeat_region   13038. .13317
                     /rpt_family="Alu"
     repeat_region   13680. .13793
                     /rpt_family="MIR"
     repeat_region   13814. .14096
                     /rpt_family="Alu"
     repeat_region   14189. .15007
                     /rpt_family="L1"
     repeat_region   15050. .15128
                     /rpt_family="ERVL"
     repeat_region   15163. .15327
                     /rpt_family="ERVL"
     repeat_region   16006. .16166
                     /rpt_family="AcHobo"
     repeat_region   16167. .16471
                     /rpt_family="Alu"
     repeat_region   16564. .16913
                     /rpt_family="Alu"
     repeat_region   17764. .17911
                     /rpt_family="Alu"
     repeat_region   17926. .18141
                     /rpt_family="L1"
     repeat_region   18161. .18681
                     /rpt_family="MaLR"
     repeat_region   18902. .18984
                     /rpt_family="GA-rich"
     repeat_region   19076. .19203
                     /rpt_family="L1"

  Query Match             1.7%;  Score 53;  DB 8;  Length 134184;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37245 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 37193


RESULT 27
BS000110
LOCUS       BS000110              157686 bp    DNA     linear   PRI 12-JUN-2004
DEFINITION  Pan troglodytes chromosome 22 clone:PTB-090P24, map 22, complete
            sequences.
ACCESSION   BS000110 BA000046
VERSION     BS000110.1  GI:37537377
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   The International Chimpanzee Chromosome 22 Consortium.
  TITLE     DNA sequence and comparative analysis of chimpanzee chromosome 22
  JOURNAL   Nature 429, 382-388 (2004)
REFERENCE   2  (bases 1 to 157686)
  AUTHORS   Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,
            Fujiyama,A. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            ------------- Genome Center
            Center: RIKEN Genomic Sciences Center
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            ------------- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: PTB-090P24
            ------------- Summary Statistics
            Sequencing vector: pUC18,pUC13,pTZ19R; 100% of reads Chemistry:
            Dye-terminator Big Dye and ET; 100% of reads Assembly program:
            Phrap; version 0.990329
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            -------------
            Source information:
            The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/mchimp43.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR:  pBACe3.6
            The CHORI-251 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/chimpanzee251.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR:  pTARBAC2.1
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS145
            The PTF22 chimpanzee Fosmid library was prepared from DNA isolated
            from cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS143
            -------------
            Sequence Quality Assessment:
            This entry has been annotated with sequence
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than 1 error in
            10,000 bp.
            ----------------------
            Neighboring clones: RP43-150L08(left) and PTB-093K04(right).
FEATURES             Location/Qualifiers
     source          1. .157686
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:9598"
                     /chromosome="22"
                     /clone="PTB-090P24"
                     /clone_lib="PTB1 chimpanzee BAC"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 157686;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123811 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 123863


RESULT 28
AL645940
LOCUS       AL645940              158033 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-157A10 on chromosome 6 contains
            the 3' end of the gene for a novel protein similar to major
            histocompatibility complex, class II, DP beta 2 pseudogene, a major
            histocompatibility complex, class II, DP beta 1 pseudogene, two
            putative novel transcripts, the COL11A2 gene for collagen, type XI,
            alpha 2, the RXRB gene for retinoid X receptor, beta, the HKE4 gene
            for HLA-class II region expressed gene KE4, the FABGL gene for
            hydroxysteroid (17-beta) dehydrogenase 8, the RING1 gene for ring
            finger protein 1, a zinger finger protein pseudogene, a HIV TAT
            specific factor 1 (HTATSF1) pseudogene, a novel transcript, the
            SACM2L gene for SAC2 suppressor of actin mutations 2-like (yeast),
            the 5' end of the RPS18 gene for ribosomal protein S18 and six CpG
            islands, complete sequence.
ACCESSION   AL645940
VERSION     AL645940.4  GI:18564736
KEYWORDS    HTG; ARE-1; ARE1; class II region expressed gene KE4; COL11A2;
            collagen, type XI, alpha 2; CpG island; D6S115E; D6S2244E;
            D6S2245E; DFNA13; FABG; FABGL; H-2RIIBP; H2-KE4; H2-KE6; HIV TAT
            specific factor 1; HKE3; HKE4; HKE5; HKE6; HLA-DPB2; HSD17B8;
            HTATSF1; hydroxysteroid dehydrogenase 8; KE-3; KE3; KE4; KE6; major
            histocompatibility complex, class II; major histocompatibility
            complex, class II, DP bet; NR2B2; PARP; RCoR-1; retinoid X
            receptor, beta; ribosomal protein S18; ring finger protein 1;
            RING1; RING2; RING5; RNF1; RP5-1033B10; RPS18; RXRB; SAC2; SAC2
            suppressor of actin mutations 2-like; SACM2L; STL3; TAT-SF1; zinc
            finger protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 158033)
  AUTHORS   Leongamornlert,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 6, 2002 this sequence version replaced gi:18070925.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-157A10 is
            from a CHORI-501 human bac - PGF cell line library VECTOR:
            pTARBAC2.1
            This sequence was generated from part of bacterial clone contigs
            constructed by the MHC Haplotype Consortium and collaborators.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .158033
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-157A10"
                     /clone_lib="CHORI-501"
     misc_feature    2000
                     /note="Clone_right_end: XXbac-138A21"
     gene            join(AL645931.7:120452. .120551,AL645931.7:124893. .124899,
                     2001. .2257,12865. .13125,13671. .13793,14102. .14143)
                     /locus_tag="XXbac-BPG157A10.5-001"
                     /pseudo
     CDS             join(AL645931.7:120452. .120551,AL645931.7:124893. .124899,
                     2001. .2257,12865. .13125,13671. .13793,14102. .14143)
                     /locus_tag="XXbac-BPG157A10.5-001"
                     /note="match: proteins: O19702 P01916 P13763 Q30166 Q31163
                     Q31633 Q95HC1 Q95IH2"
                     /pseudo
                     /codon_start=1
                     /product="major histocompatibility complex, class II, DP
                     beta 2 (pseudogene)"
     gene            join(AL645931.7:120387. .120551,AL645931.7:124893. .124899,
                     2001. .2257,3428. .4186)
                     /locus_tag="XXbac-BPG157A10.4-001"
     mRNA            join(AL645931.7:120387. .120551,AL645931.7:124893. .124899,
                     2001. .2257,3428. .4186)
                     /locus_tag="XXbac-BPG157A10.4-001"
                     /note="match: ESTs: AI971689.1 BG534679.1
                     match: cDNAs: BC017967.1"
     gene            complement(16239. .16389)
                     /locus_tag="XXbac-BPG157A10.3-001"
                     /pseudo
     CDS             complement(16239. .16389)
                     /locus_tag="XXbac-BPG157A10.3-001"
                     /note="match: proteins: P20036 Q8SP47 Q9TPX3 Q9XS10"
                     /pseudo
                     /codon_start=1
                     /product="major histocompatibility complex, class II, DP
                     alpha 2 (pseudogene)"
     gene            join(29820. .29983,32252. .32789)
                     /locus_tag="XXbac-BPG157A10.15-001"
     mRNA            join(29820. .29983,32252. .32789)
                     /locus_tag="XXbac-BPG157A10.15-001"
                     /note="match: ESTs: BG033647.1"
     gene            complement(47731. .77536)
                     /gene="COL11A2"
                     /locus_tag="XXbac-BPG157A10.2-001"
     mRNA            complement(join(47731. .48129,48779. .48855,49304. .49510,
                     49889. .50001,50586. .50853,50968. .51021,51208. .51243,
                     51550. .51603,51757. .51864,52107. .52160,52301. .52354,
                     52463. .52570,52836. .52889,53556. .53609,53743. .53796,
                     53992. .54045,54420. .54527,54868. .54921,55084. .55137,
                     55363. .55416,55594. .55647,55847. .55954,56151. .56204,
                     56308. .56361,56504. .56611,56750. .56857,57087. .57140,
                     57326. .57415,57567. .57674,58097. .58150,58385. .58438,
                     58539. .58592,58738. .58782,58910. .58963,59048. .59092,
                     59202. .59255,59564. .59617,60611. .60718,61053. .61106,
                     61296. .61340,61470. .61523,61758. .61802,62029. .62082,
                     62218. .62262,62449. .62502,62699. .62743,63169. .63222,
                     63345. .63389,63472. .63525,63716. .63769,63957. .64010,
                     64110. .64163,64268. .64321,64468. .64524,64756. .64842,
                     65295. .65369,65705. .65767,66001. .66042,66148. .66207,
                     70029. .70091,71664. .71855,73399. .73561,74015. .74225,
                     74357. .74506,77196. .77536))
```

```
                     /gene="COL11A2"
                     /locus_tag="XXbac-BPG157A10.2-001"
                     /product="collagen type XI alpha 2"
                     /note="match: ESTs: AA614321.1 AU145122.1 BQ181743.1
                     BQ182028.1
                     match: cDNAs: D38412.1 J04974.1 L18987.1 U16789.1"
     CDS             complement(join(47937. .48129,48779. .48855,49304. .49510,
                     49889. .50001,50586. .50853,50968. .51021,51208. .51243,
                     51550. .51603,51757. .51864,52107. .52160,52301. .52354,
                     52463. .52570,52836. .52889,53556. .53609,53743. .53796,
                     53992. .54045,54420. .54527,54868. .54921,55084. .55137,
                     55363. .55416,55594. .55647,55847. .55954,56151. .56204,
                     56308. .56361,56504. .56611,56750. .56857,57087. .57140,
                     57326. .57415,57567. .57674,58097. .58150,58385. .58438,
                     58539. .58592,58738. .58782,58910. .58963,59048. .59092,
                     59202. .59255,59564. .59617,60611. .60718,61053. .61106,
                     61296. .61340,61470. .61523,61758. .61802,62029. .62082,
                     62218. .62262,62449. .62502,62699. .62743,63169. .63222,
                     63345. .63389,63472. .63525,63716. .63769,63957. .64010,
                     64110. .64163,64268. .64321,64468. .64524,64756. .64842,
                     65295. .65369,65705. .65767,66001. .66042,66148. .66207,
                     70029. .70091,71664. .71855,73399. .73561,74015. .74225,
                     74357. .74506,77196. .77277))
                     /gene="COL11A2"
                     /locus_tag="XXbac-BPG157A10.2-001"
                     /standard_name="OTTHUMP00000029040"
                     /note="match: proteins: AAC17465 P13942 Q64739"
                     /codon_start=1
                     /product="collagen type XI alpha 2"
                     /protein_id="CAI18063.1"
                     /db_xref="GI:55961281"
                     /db_xref="GOA:Q5SUI8"
                     /db_xref="InterPro:IPR000885"
                     /db_xref="InterPro:IPR001791"
                     /db_xref="InterPro:IPR003129"
                     /db_xref="InterPro:IPR008160"
                     /db_xref="InterPro:IPR008161"
                     /db_xref="UniProt/TrEMBL:Q5SUI8"
                     /translation="MERCSRCHRLLLLLPLVLGLSAAPGWAGAPPVDVLRALRFPSLP
                     DGVRRAKGICPADVAYRVARPAQLSAPTRQLFPGGFPKDFSLLTVVRTRPGLQAPLLT
                     LYSAQGVRQLGLELGRPVRFLYEDQTGRPQPPSQPVFRGLSLADGKWHRVAVAVKGQS
                     VTLIVDCKKRVTRPLPRSARPVLDTHGVIIFGARILDEEVFEGDVQELAIVPGVQAAY
                     ESCEQKELECEGGQRERPQNQQPHRAQRSPQQQPSRLHRPQNQEPQSQDPTPGEEEEI
                     LESSLLPPLEEAAHGPRGLKGEKGEPAVLEPGMLVEGPPGPEGPAGLIGPPGIQGNPG
                     PVGDPGERGPPGRAGLPGSDGAPGPPGTSLMLPFRFGSGGGDKGPVVAAQEAQAQAIL
                     QQARLALRGPPGPMGYTGRPGPLGQPGSPGLKGESGDLGPQGPRGPQGLTGPPGKAGR
                     RGRAGADGARGMPGDPGVKGDRGFDGLPGLPGEKGHRGDTGAQGLPGPPGEDGERGDD
                     GEIGPRGLPGESGPRGLLGPKGPPGIPGPPGVRGMDGPQGPKGSLGPQGEPGPPGQQG
                     TPGTQGLPGPQGAIGPHGEKGPQGKPGLPGMPGSDGPPGHPGKEGPPGTKGNQGPSGP
                     QGPLGYPGPRGVKGVDGIRGLKGHKGEKGEDGFPGFKGDIGVKGDRGEVGVPGSRGED
                     GPEGPKGRTGPTGDPGPPGLMGEKGKLGVPGLPGYPGRQGPKGSLGFPGFPGASGEKG
                     ARGLSGKSGPRGERGPTGPRGQRGPRGATGKSGAKGTSGGDGPHGPPGERGLPGPQGP
                     NGFPGPKGPPGPPGKDGLPGHPGQRGEVGFQGKTGPPGPPGVVGPQGAAGETGPMGER
                     GHPGPPGPPGEQGLPGTAGKEGTKGDPGPPGAPGKDGPAGLRGFPGERGLPGTAGGPG
                     LKGNEGPSGPPGPAGSPGERGAAGSGGPIGPPGRPGPQGPPGAAGEKGVPGEKGPIGP
                     TGRDGVQGPVGLPGPAGPPGVAGEDGDKGEVGDPGQKGTKGNKGEHGPPGPPGPIGPV
                     GQPGAAGADGEPGARGPQGHPGAKGDEGTRGFNGPPGPIGLQGLPGPSGEKGETGDVG
                     PMGPPGPPGPRGPAGPNGADGPQGPPGGVGNLGPPGEKGEPGESGSPGIQGEPGVKGP
                     RGERGEKGESGQPGEPGPPGPKGPTGDDGPKGNPGPVGFPGDPGPPGEGGPRGQDGAK
                     GDRGEDGEPGQPGSPGPTGENGPPGPLGKRGPAGSPGSEGRQGGKGAKGDPGAIGAPG
                     KTGPVGPAGPAGKPGPDGLRGLPGSVGQQGRPGATGQAGPPGPVGPPGLPGLRGDAGA
                     KGEKGHPGLIGLIGPPGEQGEKGDRGLPGPQGSPGQKGEMGIPGASGPIGPGGPPGLP
                     GPAGPKGAKGATGPGGPKGEKGVQGPPGHPGPPGEVIQPLPIQMPKKTRRSVDGSRLM
                     QEDEAIPTGGAPGSPGGLEEIFGSLDSLRREEIEQMRRPTGTQDSPARTCQDLKLCHPE
                     LPDGEYWVDPNQGCARDAFRVFCNFTAGGETCVTPRDDVTQFSYVDSEGSPVGVVQLT
                     FLRLLSVSAHQDVSYPCSGAARDGPLRLRGANEDELSPETSPYVKEFRDGCQTQQGRT
                     VLEVRTPVLEQLPVLDASFSDSARVGLSNPQVLAAWPMCPLSLPLGLGLAGTACYLRA
                     RMLPQGRNGIVWGVCRAASRRWLEY"
        gene         complement(join(78625. .79866,79982. .80087,80400. .80491,
                     80607. .80739,80944. .81073,81471. .81643,82799. .82978,
                     83345. .83501,84206. .84453,85279. .85699))

Query Match          1.7%;  Score 53;  DB 8;  Length 158033;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123440 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 123492


RESULT 29
AC022960
LOCUS       AC022960              171051 bp    DNA     linear   PRI 20-FEB-2002
DEFINITION  Homo sapiens chromosome 18, clone RP11-210K20, complete sequence.
ACCESSION   AC022960
VERSION     AC022960.13  GI:18767551
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171051)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-210K20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171051)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 171051)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 171051)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
```

```
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 20, 2002 this sequence version replaced gi:18693515.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6420
                Center clone name: 210_K_20
            --------------.
FEATURES             Location/Qualifiers
     source          1. .171051
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-210K20"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   1. .176
                     /rpt_family="MIR"
     repeat_region   223. .251
                     /rpt_family="(TG)n"
     repeat_region   537. .643
                     /rpt_family="MER5A"
     repeat_region   1820. .2239
                     /rpt_family="L1MA7"
     repeat_region   2240. .2530
                     /rpt_family="AluJb"
     repeat_region   2531. .4997
                     /rpt_family="L1MA7"
     repeat_region   4998. .5294
                     /rpt_family="AluYa5"
     repeat_region   5295. .5770
                     /rpt_family="L1MA7"
     repeat_region   complement(5840. .5966)
                     /rpt_family="FLAM_C"
     repeat_region   7146. .7178
                     /rpt_family="AT_rich"
     repeat_region   7335. .7362
                     /rpt_family="(TTTA)n"
     repeat_region   complement(7365. .7646)
                     /rpt_family="AluJo"
     repeat_region   complement(7678. .7966)
                     /rpt_family="AluSx"
     repeat_region   8635. .9181
                     /rpt_family="LTR10A"
     repeat_region   complement(9652. .9786)
                     /rpt_family="MER5B"
     repeat_region   complement(9842. .10444)
                     /rpt_family="LTR9B"
     repeat_region   complement(10448. .10690)
                     /rpt_family="HUERS-P3"
     repeat_region   10727. .11026
                     /rpt_family="AluSx"
     repeat_region   complement(11319. .11430)
                     /rpt_family="LTR9B-int"
     repeat_region   complement(11427. .12609)
                     /rpt_family="LTR9B-int"
     repeat_region   12610. .12919
                     /rpt_family="AluSp"
     repeat_region   complement(12920. .13216)
                     /rpt_family="LTR9B-int"
     repeat_region   13217. .13527
                     /rpt_family="AluSq"
     repeat_region   complement(13528. .13969)
                     /rpt_family="LTR9B-int"
     repeat_region   complement(13972. .14575)
                     /rpt_family="LTR9B"
     repeat_region   14583. .15192
                     /rpt_family="MER4B"
     repeat_region   complement(15330. .15674)
                     /rpt_family="MER21B"
     repeat_region   complement(15675. .15989)
                     /rpt_family="AluJo"
     repeat_region   complement(15990. .16412)
                     /rpt_family="MER21B"
     repeat_region   complement(16412. .16531)
                     /rpt_family="MER21B"
     repeat_region   complement(17688. .19067)
                     /rpt_family="L1P"
     repeat_region   19061. .19138
                     /rpt_family="L1PA13"
     repeat_region   19139. .19201
                     /rpt_family="L1PA13"
     repeat_region   19202. .19483
                     /rpt_family="AluSq"
     repeat_region   19484. .19882
                     /rpt_family="L1PA13"
     repeat_region   21669. .22281
                     /rpt_family="L1MA8"
     repeat_region   complement(22282. .22696)
                     /rpt_family="Tigger2a"
     repeat_region   22697. .22886
                     /rpt_family="L1MA8"
     repeat_region   22888. .22908
                     /rpt_family="AT_rich"
     repeat_region   complement(23103. .23255)
                     /rpt_family="MER110"
     repeat_region   23290. .23427
                     /rpt_family="(TA)n"
     repeat_region   complement(23697. .24076)
                     /rpt_family="MSTC"
     repeat_region   complement(25025. .25187)
                     /rpt_family="MIR"
     repeat_region   complement(25208. .25502)
                     /rpt_family="AluSg"
     repeat_region   25814. .25847
                     /rpt_family="(TTTA)n"
     repeat_region   complement(26782. .27105)

  Query Match             1.7%;  Score 53;  DB 8;  Length 171051;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90385 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 90437
```

```
RESULT 30
HS1033B10/c
LOCUS       HS1033B10              175737 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP5-1033B10 on chromosome
            6p21.2-21.31 Contains the 3' end of the C6ORF11 gene for chromosome
            6 open reading frame 11 (BING4), the B3GALT4 gene for
            UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase polypeptide 4
            (GalT4,Gal-T2,beta3GALT4,beta3Gal-T4), the RPS18 gene for 40S
            ribosomal protein S18 (KE3,HKE3,KE-3), the SACM2L gene for SAC2
            suppressor of actin mutations 2-like
            (yeast)(ARE1,SAC2,ARE-1,RP5-1033B10), a pseudogene similar to part
            of HTATSF1 (HIV TAT specific factor 1), the RING1 gene for ring
            finger protein 1 (RNF1), the HSD17B8 gene for hydroxysteroid
            (17-beta) dehydrogenase 8
            (KE6,FABG,HKE6,FABGL,RING2,H2-KE6,D6S2245E), the HKE4 gene for HLA
            class II region protein KE4 (KE4,RING5,H2-KE4,D6S115E,D6S2244E),
            the RXRB gene for retinoid X receptor beta (NR2B2,RCoR-1,H-2RIIBP),
            the COL11A2 gene for collagen type XI alpha 2
            (HKE5,PARP,STL3,DFNA13), an HLA class II histocompatability antigen
            D or S beta pseudogene,the gene for a novel protein similar to
            major histocompatibility complex class II DP beta 1 precursor,
            three novel genes and seven CpG islands, complete sequence.
ACCESSION   AL031228
VERSION     AL031228.1  GI:3646023
KEYWORDS    HTG; 2RIIBP; ARE-1; ARE1; B3GALT4; beta3Gal-T4; beta3GALT4; BING4;
            C6ORF11; COL11A2; collagen; CpG island; D6S115E; D6S2244E;
            D6S2245E; DFNA13; FABG; FABGL; GalT4; Gal_T2; H2-KE4; H2-KE6; HKE3;
            HKE4; HKE5; HKE6; HLA_DPB1; HSD17B8; HTATSF1; KE-3; KE3; KE4; KE6;
            NR2B2; PARP; RCoR-1; retinoid X receptor beta; ribosomal protein;
            ring finger; RING1; RING2; RING5; RNF1; RPS18; RXRB; SAC2; SACM2L;
            STL3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 175737)
  AUTHORS   Tubby,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 24, 1998 this sequence version replaced gi:3550015.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP5-1033B10 is from the library RPCI-5 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .175737
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704B101033"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p21.2-21.31"
                     /clone="RP5-1033B10"
                     /clone_lib="RPCI-5"
     gene            join(Z97184.1:31787. .31978,Z97184.1:32067. .32276,
                     Z97184.1:32386. .32466,Z97184.1:32610. .32722,
                     Z97184.1:32845. .32932,Z97184.1:33054. .33115,
                     Z97184.1:33362. .33466,Z97184.1:33575. .33725,
                     Z97184.1:33854. .33989,Z97184.1:34186. .34285,5240. .5553,
                     5705. .5799,6368. .6463,6617. .6730,6853. .7119)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-001"
     mRNA            join(Z97184.1:31787. .31978,Z97184.1:32067. .32276,
                     Z97184.1:32386. .32466,Z97184.1:32610. .32722,
                     Z97184.1:32845. .32932,Z97184.1:33054. .33115,
                     Z97184.1:33362. .33466,Z97184.1:33575. .33725,
                     Z97184.1:33854. .33989,Z97184.1:34186. .34285,5240. .5553,
                     5705. .5799,6368. .6463,6617. .6730,6853. .7119)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-001"
                     /product="chromosome 6 open reading frame 11"
                     /note="match: ESTs: AA118452.1 AA156666.1 AA185515.1
                     AI225748.1 AI538411.1 AI654175.1 AI670118.1 AI745012.1
                     AI745013.1 AI828707.1 AI867295.1 AI968094.1 AI990969.1
                     AL526165.1 AL526244.1 AL527886.1 AL545883.1 AL552140.1
                     AL554794.1 AL556378.1 AL556379.1 AL557009.1 AL558638.1
                     AL561268.1 AL571437.1 AL575523.1 AL577919.1 AL579740.1
                     AL579741.1 AL580412.1 AL581355.1 AL582543.1 AL582963.1
                     AL583014.1 AL602879.1 AL708305.1 AV605813.1 AV611266.1
                     AV611267.1 AW192492.1 AW250879.1 AW307819.1 AW416899.1
                     AW445475.1 AW515224.1 BE234338.1 BE243105.1 BE254716.1
                     BE388763.1 BE407443.1 BE536575.1 BE617010.1 BE619296.1
                     BE622746.1 BE676236.1 BE797604.1 BE888869.1 BE890194.1
                     BE890448.1 BE892221.1 BE966732.2 BF035791.1 BF182424.1
                     BF236106.1 BF312238.1 BF312319.1 BF313570.1 BF345554.1
                     BF584138.1 BF781258.1 BF889575.1 BF971752.1 BF974816.1
                     BG008501.1 BG010013.1 BG034349.1 BG115980.1 BG119460.1
                     BG250855.1 BG283791.1 BG476159.1 BG491447.1 BG491645.1
                     BG491994.1 BG684964.1 BG690303.1 BG702426.1 BG752850.1
                     BG756009.1 BG765705.1 BI196137.1 BI196220.1 BI261976.1
                     BI338173.1 BI345792.1 BI415032.1 BI524736.1 BI562558.1
                     BI604022.1 BI759414.1 BI771500.1 BI828412.1 BI909253.1
                     BI910702.1 BM006299.1 BM007199.1 BM008019.1 BM010702.1
                     BM107347.1 BM471921.1 BM475369.1 BM510626.1 BM560608.1
                     BM920979.1 BM923020.1 BM923858.1 BM970926.1 BQ005630.1
                     BQ014292.1 BQ060007.1 BQ061145.1 BQ062625.1 BQ068624.1
                     BQ073118.1 BQ436175.1 H09026.1 H52171.1 H52575.1 T97040.1
                     match: cDNAs: BC000388.1 BC013527.1 BC026431.1"
     CDS             join(Z97184.1:31910. .31978,Z97184.1:32067. .32276,
                     Z97184.1:32386. .32466,Z97184.1:32610. .32722,
                     Z97184.1:32845. .32932,Z97184.1:33054. .33115,
                     Z97184.1:33362. .33466,Z97184.1:33575. .33725,
                     Z97184.1:33854. .33989,Z97184.1:34186. .34285,5240. .5553,
                     5705. .5799,6368. .6463,6617. .6730,6853. .6951)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-001"
                     /standard_name="OTTHUMP00000017904"
                     /note="match: proteins: O15213 Q96KL3 Q9Z0H1"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 11"
                     /protein_id="CAA20229.1"
                     /db_xref="GI:3820978"
                     /translation="METAPKPGKDVPPKKDKLQTKRKKPRRYWEEETVPTTAGASPGP
                     PRNKKNRELRPQRPKNAYILKKSRISKKPQVPKKPREWKNPESQRGLSGAQDPFPGPA
                     PVPVEVVQKFCRIDKSRKLPHSKAKTRSRLEVAEAEEEETSIKAARSELLLAEEPGFL
                     EGEDGEDTAKICQADIVEAVDIASAAKHFDLNLRQFGPYRLNYSRTGRHLAFGGRRGH
                     VAALDWVTKKLMCEINVMEAVRDIRFLHSEALLAVAQNRWLHIYDNQGIELHCIRRCD
                     RVTRLEFLPFHFLLATASETGFLTYLDVSVGKIVAALNARAGRLDVMSQNPYNAVIHL
                     GHSNGTVSLWSPAMKEPLAKILCHRGGVRAVAVDSTGTYMATSGLDHQLKIFDLRGTY
                     QPLSTRTLPHGAGHLAFSQRGLLVAGMGDVVNIWAGQGKASPPSLEQPYLTHRLSGPV
```

```
                     HGLQFCPFEDVLGVGHTGGITSMLVPGAGEPNFDGLESNPYRSRKQRQEWEVKALLEK
                     VPAELICLDPRALAEVDVISLEQGKKEQIERLGYDPQAKAPFQPKPKQKGRSSTASLV
                     KRKRKVMDEEHRDKVRQSLQQQHHKEAKAKPTGARPSALDRFVR"
     gene            join(Z97184.1:33855. .33989,Z97184.1:34186. .34285,
                     5240. .5799,6368. .6463,6617. .6730,6853. .6912)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-002"
     mRNA            join(Z97184.1:33855. .33989,Z97184.1:34186. .34285,
                     5240. .5799,6368. .6463,6617. .6730,6853. .6912)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-002"
                     /product="chromosome 6 open reading frame 11"
                     /note="match: ESTs: BE164701.1 BE667540.1 BG697282.1
                     BG912913.1"
     gene            join(Z97184.1:31873. .31978,Z97184.1:32067. .32188,
                     Z97184.1:33575. .33725,Z97184.1:33854. .33989,
                     Z97184.1:34186. .34285,5240. .5519)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-004"
     mRNA            join(Z97184.1:31873. .31978,Z97184.1:32067. .32188,
                     Z97184.1:33575. .33725,Z97184.1:33854. .33989,
                     Z97184.1:34186. .34285,5240. .5519)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-004"
                     /product="chromosome 6 open reading frame 11"
                     /note="match: ESTs: BG470714.1"
     gene            join(5316. .5412,5705. .5799,6368. .6463,6617. .6730,
                     6853. .7119)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-007"
     mRNA            join(5316. .5412,5705. .5799,6368. .6463,6617. .6730,
                     6853. .7119)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-007"
                     /product="chromosome 6 open reading frame 11"
                     /note="match: ESTs: BF476741.1"
     gene            join(5374. .5553,5705. .5799,6368. .6463,6617. .7122)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-006"
     mRNA            join(5374. .5553,5705. .5799,6368. .6463,6617. .7122)
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-006"
                     /product="chromosome 6 open reading frame 11"
                     /note="match: ESTs: AA599154.1 AA627958.1 AI204504.1
                     AI524223.1 AW193506.1 AW294189.1 BE504644.1 BE617532.1
                     H07040.1 R39981.1 R40596.1"
     polyA_signal    7094. .7099
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-006"
     polyA_site      7111
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-006"
     polyA_site      7119
                     /gene="C6orf11"
                     /locus_tag="RP5-1033B10.11-006"
     polyA_site      7122
                     /gene="C6orf11"

  Query Match             1.7%;  Score 53;  DB 8;  Length 175737;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47829 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 47777


RESULT 31
AC021102
LOCUS       AC021102              179462 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-129G15, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC021102
VERSION     AC021102.3  GI:8567961
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 179462)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 179462)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 15, 2000 this sequence version replaced gi:7230893.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0129G15
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 75%
            Sequencing vector: plasmid; 25%
            Chemistry: Dye-primer ET; 75% of reads
            Chemistry: Dye-terminator Big Dye; 25% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 165207 bases at least Q40
            Consensus quality: 169791 bases at least Q30
            Consensus quality: 172552 bases at least Q20
            Insert size: 177000; agarose-fp
            Insert size: 177062; sum-of-contigs
            Quality coverage: 3.62 in Q20 bases; agarose-fp
            Quality coverage: 3.66 in Q20 bases; sum-of-contigs
            -------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1892: contig of 1892 bp in length
            *     1893     1992: gap of unknown length
            *     1993     4343: contig of 2351 bp in length
            *     4344     4443: gap of unknown length
            *     4444     6559: contig of 2116 bp in length
            *     6560     6659: gap of unknown length
            *     6660    10517: contig of 3858 bp in length
            *    10518    10617: gap of unknown length
            *    10618    14073: contig of 3456 bp in length
            *    14074    14173: gap of unknown length
            *    14174    16899: contig of 2726 bp in length
            *    16900    16999: gap of unknown length
            *    17000    20655: contig of 3656 bp in length
            *    20656    20755: gap of unknown length
            *    20756    24533: contig of 3778 bp in length
            *    24534    24633: gap of unknown length
            *    24634    27750: contig of 3117 bp in length
            *    27751    27850: gap of unknown length
            *    27851    29522: contig of 1672 bp in length
            *    29523    29622: gap of unknown length
            *    29623    33750: contig of 4128 bp in length
            *    33751    33850: gap of unknown length
            *    33851    37904: contig of 4054 bp in length
            *    37905    38004: gap of unknown length
            *    38005    43241: contig of 5237 bp in length
            *    43242    43341: gap of unknown length
```

```
            *    43342    50634: contig of 7293 bp in length
            *    50635    50734: gap of unknown length
            *    50735    59411: contig of 8677 bp in length
            *    59412    59511: gap of unknown length
            *    59512    68503: contig of 8992 bp in length
            *    68504    68603: gap of unknown length
            *    68604    79177: contig of 10574 bp in length
            *    79178    79277: gap of unknown length
            *    79278    86283: contig of 7006 bp in length
            *    86284    86383: gap of unknown length
            *    86384    96915: contig of 10532 bp in length
            *    96916    97015: gap of unknown length
            *    97016   107357: contig of 10342 bp in length
            *   107358   107457: gap of unknown length
            *   107458   119546: contig of 12089 bp in length
            *   119547   119646: gap of unknown length
            *   119647   134223: contig of 14577 bp in length
            *   134224   134323: gap of unknown length
            *   134324   147018: contig of 12695 bp in length
            *   147019   147118: gap of unknown length
            *   147119   160737: contig of 13619 bp in length
            *   160738   160837: gap of unknown length
            *   160838   179462: contig of 18625 bp in length.
FEATURES             Location/Qualifiers
     source          1. .179462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="RP11-129G15"
     misc_feature    1. .1892
                     /note="assembly_name:Contig11"
     gap             1893. .1992
                     /estimated_length=unknown
     misc_feature    1993. .4343
                     /note="assembly_name:Contig12"
     gap             4344. .4443
                     /estimated_length=unknown
     misc_feature    4444. .6559
                     /note="assembly_name:Contig13"
     gap             6560. .6659
                     /estimated_length=unknown
     misc_feature    6660. .10517
                     /note="assembly_name:Contig14"
     gap             10518. .10617
                     /estimated_length=unknown
     misc_feature    10618. .14073
                     /note="assembly_name:Contig15"
     gap             14074. .14173
                     /estimated_length=unknown
     misc_feature    14174. .16899
                     /note="assembly_name:Contig16"
     gap             16900. .16999
                     /estimated_length=unknown
     misc_feature    17000. .20655
                     /note="assembly_name:Contig17"
     gap             20656. .20755
                     /estimated_length=unknown
     misc_feature    20756. .24533
                     /note="assembly_name:Contig18"
     gap             24534. .24633
                     /estimated_length=unknown
     misc_feature    24634. .27750
                     /note="assembly_name:Contig19"
     gap             27751. .27850
                     /estimated_length=unknown
     misc_feature    27851. .29522
                     /note="assembly_name:Contig20
                     clone_end:SP6
                     vector_side:left"
     gap             29523. .29622
                     /estimated_length=unknown
     misc_feature    29623. .33750
                     /note="assembly_name:Contig21"
     gap             33751. .33850
                     /estimated_length=unknown
     misc_feature    33851. .37904
                     /note="assembly_name:Contig22
                     clone_end:T7
                     vector_side:right"
     gap             37905. .38004
                     /estimated_length=unknown
     misc_feature    38005. .43241
                     /note="assembly_name:Contig23"
     gap             43242. .43341
                     /estimated_length=unknown
     misc_feature    43342. .50634
                     /note="assembly_name:Contig24"
     gap             50635. .50734
                     /estimated_length=unknown
     misc_feature    50735. .59411
                     /note="assembly_name:Contig25"
     gap             59412. .59511
                     /estimated_length=unknown
     misc_feature    59512. .68503
                     /note="assembly_name:Contig26"
     gap             68504. .68603
                     /estimated_length=unknown
     misc_feature    68604. .79177
                     /note="assembly_name:Contig27"
     gap             79178. .79277
                     /estimated_length=unknown
     misc_feature    79278. .86283
                     /note="assembly_name:Contig28"
     gap             86284. .86383
                     /estimated_length=unknown
     misc_feature    86384. .96915
                     /note="assembly_name:Contig29"
     gap             96916. .97015
                     /estimated_length=unknown
     misc_feature    97016. .107357
                     /note="assembly_name:Contig30"
     gap             107358. .107457
                     /estimated_length=unknown
     misc_feature    107458. .119546
                     /note="assembly_name:Contig31"
     gap             119547. .119646
                     /estimated_length=unknown
     misc_feature    119647. .134223
                     /note="assembly_name:Contig32"
     gap             134224. .134323
                     /estimated_length=unknown
     misc_feature    134324. .147018

  Query Match             1.7%;  Score 53;  DB 14;  Length 179462;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74903 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 74955


RESULT 32
AC007332
LOCUS       AC007332              185000 bp    DNA     linear   HTG 14-DEC-1999
DEFINITION  Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***, 14
            unordered pieces.
ACCESSION   AC007332
VERSION     AC007332.11  GI:6573831
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 185000)
  AUTHORS   Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185000)
  AUTHORS   Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Department of Genetics, Stanford Human
            Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT     On Dec 14, 1999 this sequence version replaced gi:6137906.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1233: contig of 1233 bp in length
            *    1234     1994: gap of unknown length
            *    1995     3971: contig of 1977 bp in length
            *    3972     4732: gap of unknown length
            *    4733     6959: contig of 2227 bp in length
            *    6960     7720: gap of unknown length
            *    7721     9753: contig of 2033 bp in length
            *    9754    10514: gap of unknown length
            *   10515    12313: contig of 1799 bp in length
            *   12314    13074: gap of unknown length
            *   13075    15817: contig of 2743 bp in length
            *   15818    16578: gap of unknown length
            *   16579    21924: contig of 5346 bp in length
            *   21925    22685: gap of unknown length
            *   22686    27521: contig of 4836 bp in length
            *   27522    28282: gap of unknown length
            *   28283    33545: contig of 5263 bp in length
            *   33546    34306: gap of unknown length
            *   34307    51372: contig of 17066 bp in length
            *   51373    52133: gap of unknown length
            *   52134    73810: contig of 21677 bp in length
            *   73811    74570: gap of unknown length
            *   74571   101382: contig of 26812 bp in length
            *  101383   102142: gap of unknown length
            *  102143   143256: contig of 41114 bp in length
            *  143257   144016: gap of unknown length
            *  144017   185000: contig of 40984 bp in length.
FEATURES             Location/Qualifiers
     source          1. .185000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
     gap             1234. .1994
                     /estimated_length=unknown
     gap             3972. .4732
                     /estimated_length=unknown
     gap             6960. .7720
                     /estimated_length=unknown
     gap             9754. .10514
                     /estimated_length=unknown
     gap             12314. .13074
                     /estimated_length=unknown
     gap             15818. .16578
                     /estimated_length=unknown
     gap             21925. .22685
                     /estimated_length=unknown
     gap             27522. .28282
                     /estimated_length=unknown
     gap             33546. .34306
                     /estimated_length=unknown
     gap             51373. .52133
                     /estimated_length=unknown
     gap             73811. .74570
                     /estimated_length=unknown
     gap             101383. .102142
                     /estimated_length=unknown
     gap             143257. .144016
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 185000;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168527 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 168579


RESULT 33
AP001333
LOCUS       AP001333               189181 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens chromosome 18 clone RP11-675J18 map 18q12, WORKING
            DRAFT SEQUENCE, 25 unordered pieces.
ACCESSION   AP001333
VERSION     AP001333.3  GI:9186910
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 189181)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 189,181 genomic DNA of 18q12
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 189181)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-42-778-9923, Fax:81-42-778-9924)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:8117256.
            -------------- Genome Center
            Center: RIKEN Genomic Sciences Center(GSC)
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
            Center project name: HumDraft18
            Center clone name: RP11-675J18
            -------------- Summary Statistics
            Sequencing vector: PCR products; 100% of reads
            Chemistry: Dye-terminator ET-amersham; 100% of reads Assembly
            program: Phrap; version 0.990329
            Consensus quality: 181064 bases at least Q40
            Consensus quality: 183725 bases at least Q30
            Consensus quality: 185304 bases at least Q20
            Insert size: 186781; sum-of-contigs
            Quality coverage: 9.17x in Q20 bases; sum-of-contigs
            --------------.
            NOTE: This is a 'working draft' sequence. It currently consists of
            25 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
            1     37335 contig of  37335 bp in length
            37436     67386 contig of  29951 bp in length
            67487     90245 contig of  22759 bp in length
            90346    100199 contig of   9854 bp in length
            100300    108800 contig of   8501 bp in length
```

```
108901    117127 contig of   8227 bp in length
117228    124504 contig of   7277 bp in length
124605    132198 contig of   7594 bp in length
132299    138051 contig of   5753 bp in length
138152    144498 contig of   6347 bp in length
144599    150308 contig of   5710 bp in length
150409    154983 contig of   4575 bp in length
155084    158506 contig of   3423 bp in length
158607    162373 contig of   3767 bp in length
162474    164257 contig of   1784 bp in length
164358    167268 contig of   2911 bp in length
167369    171012 contig of   3644 bp in length
171113    174768 contig of   3656 bp in length
174869    177195 contig of   2327 bp in length
177296    179508 contig of   2213 bp in length
179609    181405 contig of   1797 bp in length
181506    184223 contig of   2718 bp in length
184324    185884 contig of   1561 bp in length
185985    187859 contig of   1875 bp in length
187960    189181 contig of   1222 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*         1    37335: contig of 37335 bp in length
*     37336    37435: gap of 100 bp
*     37436    67386: contig of 29951 bp in length
*     67387    67486: gap of 100 bp
*     67487    90245: contig of 22759 bp in length
*     90246    90345: gap of 100 bp
*     90346   100199: contig of 9854 bp in length
*    100200   100299: gap of 100 bp
*    100300   108800: contig of 8501 bp in length
*    108801   108900: gap of 100 bp
*    108901   117127: contig of 8227 bp in length
*    117128   117227: gap of 100 bp
*    117228   124504: contig of 7277 bp in length
*    124505   124604: gap of 100 bp
*    124605   132198: contig of 7594 bp in length
*    132199   132298: gap of 100 bp
*    132299   138051: contig of 5753 bp in length
*    138052   138151: gap of 100 bp
*    138152   144498: contig of 6347 bp in length
*    144499   144598: gap of 100 bp
*    144599   150308: contig of 5710 bp in length
*    150309   150408: gap of 100 bp
*    150409   154983: contig of 4575 bp in length
*    154984   155083: gap of 100 bp
*    155084   158506: contig of 3423 bp in length
*    158507   158606: gap of 100 bp
*    158607   162373: contig of 3767 bp in length
*    162374   162473: gap of 100 bp
*    162474   164257: contig of 1784 bp in length
*    164258   164357: gap of 100 bp
*    164358   167268: contig of 2911 bp in length
*    167269   167368: gap of 100 bp
*    167369   171012: contig of 3644 bp in length
*    171013   171112: gap of 100 bp
*    171113   174768: contig of 3656 bp in length
*    174769   174868: gap of 100 bp
*    174869   177195: contig of 2327 bp in length
*    177196   177295: gap of 100 bp
*    177296   179508: contig of 2213 bp in length
*    179509   179608: gap of 100 bp
*    179609   181405: contig of 1797 bp in length
*    181406   181505: gap of 100 bp
*    181506   184223: contig of 2718 bp in length
*    184224   184323: gap of 100 bp
*    184324   185884: contig of 1561 bp in length
*    185885   185984: gap of 100 bp
*    185985   187859: contig of 1875 bp in length
*    187860   187959: gap of 100 bp
*    187960   189181: contig of 1222 bp in length.
FEATURES             Location/Qualifiers
     source          1. .189181
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18q12"
                     /clone="RP11-675J18"
     misc_feature    1. .37335
                     /note="assembly_fragment"
     misc_feature    37436. .67386
                     /note="assembly_fragment"
     misc_feature    67487. .90245
                     /note="assembly_fragment"
     misc_feature    90346. .100199
                     /note="assembly_fragment clone_end:T7 vector_side:left"
     misc_feature    100300. .108800
                     /note="assembly_fragment"
     misc_feature    108901. .117127
                     /note="assembly_fragment"
     misc_feature    117228. .124504
                     /note="assembly_fragment"
     misc_feature    124605. .132198
                     /note="assembly_fragment"
     misc_feature    132299. .138051
                     /note="assembly_fragment"
     misc_feature    138152. .144498
                     /note="assembly_fragment"
     misc_feature    144599. .150308
                     /note="assembly_fragment"
     misc_feature    150409. .154983
                     /note="assembly_fragment"
     misc_feature    155084. .158506
                     /note="assembly_fragment"
     misc_feature    158607. .162373
                     /note="assembly_fragment"
     misc_feature    162474. .164257
                     /note="assembly_fragment clone_end:SP6 vector_side:left"
     misc_feature    164358. .167268
                     /note="assembly_fragment"
     misc_feature    167369. .171012
                     /note="assembly_fragment"
     misc_feature    171113. .174768
                     /note="assembly_fragment"
     misc_feature    174869. .177195
                     /note="assembly_fragment"
     misc_feature    177296. .179508
                     /note="assembly_fragment"
     misc_feature    179609. .181405
                     /note="assembly_fragment"
     misc_feature    181506. .184223
                     /note="assembly_fragment"
     misc_feature    184324. .185884
                     /note="assembly_fragment"
     misc_feature    185985. .187859
                     /note="assembly_fragment"
     misc_feature    187960. .189181
                     /note="assembly_fragment"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 189181;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137657 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 137709
```

```
RESULT 34
AC120502
LOCUS       AC120502              190092 bp    DNA     linear   HTG 07-MAY-2002
DEFINITION  Lemur catta clone LB2-287D5, WORKING DRAFT SEQUENCE, 13 unordered
            pieces.
ACCESSION   AC120502
VERSION     AC120502.1  GI:20486402
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Lemur catta (ring-tailed lemur)
  ORGANISM  Lemur catta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Strepsirhini;
            Lemuridae; Lemur.
REFERENCE   1  (bases 1 to 190092)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
            Cheng,J.-F.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 190092)
  AUTHORS   Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
            Cheng,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2002) Genome Sciences, Lawrence Berkeley National
            Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
COMMENT     Draft Sequence Produced by Berkeley PGA
            Web site: http://pga.lbl.gov
            Center Code: PGABERK
            Center Project Name: L017-287D5
            Bac Clone Name: LB2-287D5

            Additional information on comparative analysis and ordering are
            available at:
            http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=SFTPD
            Funding agent: Programs for Genomic Applications (NHLBI)
            Contact: 'Jody Schwartz' jrschwartz@lbl.gov
            if library name is LB1 to LB4, please see website
            for a description: http://www-gsd.lbl.gov/cheng/BAC.html
            Summary Statistics:
            Sequencing vector: Plasmid; pUC18
            Chemistry: Dye-terminator Big Dye
            Assembly program: Phrap version 0.990329.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2537: contig of 2537 bp in length
            *     2538     2637: gap of unknown length
            *     2638     6734: contig of 4097 bp in length
            *     6735     6834: gap of unknown length
            *     6835    14153: contig of 7319 bp in length
            *    14154    14253: gap of unknown length
            *    14254    21025: contig of 6772 bp in length
            *    21026    21125: gap of unknown length
            *    21126    33188: contig of 12063 bp in length
            *    33189    33288: gap of unknown length
            *    33289    43658: contig of 10370 bp in length
            *    43659    43758: gap of unknown length
            *    43759    59694: contig of 15936 bp in length
            *    59695    59794: gap of unknown length
            *    59795    75288: contig of 15494 bp in length
            *    75289    75388: gap of unknown length
            *    75389    93753: contig of 18365 bp in length
            *    93754    93853: gap of unknown length
            *    93854   112133: contig of 18280 bp in length
            *   112134   112233: gap of unknown length
            *   112234   136450: contig of 24217 bp in length
            *   136451   136550: gap of unknown length
            *   136551   161260: contig of 24710 bp in length
            *   161261   161360: gap of unknown length
            *   161361   190092: contig of 28732 bp in length.
FEATURES             Location/Qualifiers
     source          1. .190092
                     /organism="Lemur catta"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9447"
                     /clone="LB2-287D5"
     gap             2538. .2637
                     /estimated_length=unknown
     gap             6735. .6834
                     /estimated_length=unknown
     gap             14154. .14253
                     /estimated_length=unknown
     gap             21026. .21125
                     /estimated_length=unknown
     gap             33189. .33288
                     /estimated_length=unknown
     gap             43659. .43758
                     /estimated_length=unknown
     gap             59695. .59794
                     /estimated_length=unknown
     gap             75289. .75388
                     /estimated_length=unknown
     gap             93754. .93853
                     /estimated_length=unknown
     gap             112134. .112233
                     /estimated_length=unknown
     gap             136451. .136550
                     /estimated_length=unknown
     gap             161261. .161360
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 190092;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2970 AAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 3022
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72804 AAAAATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 72856


RESULT 35
AC092437
LOCUS       AC092437              203583 bp    DNA     linear   PRI 09-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-326I19 from 4, complete sequence.
ACCESSION   AC092437
VERSION     AC092437.4  GI:19310338
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 203583)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 203583)
  AUTHORS   Mulvaney,E. and Meyer,R.
  TITLE     The sequence of Homo sapiens BAC clone RP11-326I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 203583)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 203583)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Department of Genetics, Washington
```

```
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Mar 9, 2002 this sequence version replaced gi:15799649.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0326I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-274B16; the clone sequenced
            to the right is RP11-386I19.  Actual start of this clone is at base
            position 1 of RP11-326I19; actual end is at base position 203583 of
            RP11-326I19.

            There are single plasmid subclone regions from 9433 to 9523, 179086
            to 179114 and 195823 to 195835.  There is an unresolved
            homopolymeric run from 74097 to 74101.  Polymorphisms have been
            identified between AC025084, AC108139, AC007332 and AC092437.  Data
            from AC025084, AC108139 and  AC007332 was used to finish this
            clone, AC092437.
FEATURES                    Location/Qualifiers
     source                 1. .203583
                            /organism="Homo sapiens"
                            /mol_type="genomic DNA"
                            /db_xref="taxon:9606"
                            /chromosome="4"
                            /map="4"
                            /clone="RP11-326I19"
                            /clone_lib="RPCI-11"
     repeat_region          382. .508
                            /rpt_family="L2"
     repeat_region          522. .639
                            /rpt_family="MIR"
     repeat_region          1049. .1187
                            /rpt_family="MIR"
     repeat_region          1237. .1347
                            /rpt_family="L2"
     repeat_region          1350. .1389
                            /rpt_family="L2"
     repeat_region          1485. .1559
                            /rpt_family="MIR"
     repeat_region          1640. .2055
                            /rpt_family="MaLR"
     repeat_region          2254. .2391
                            /rpt_family="MIR"
     repeat_region          2717. .3003
                            /rpt_family="L1"
     repeat_region          3047. .3709
                            /rpt_family="MER2_type"
     repeat_region          3732. .3781
                            /rpt_family="(CATATA)n"
     repeat_region          3802. .3929
                            /rpt_family="L1"
     repeat_region          3935. .3961
                            /rpt_family="AT_rich"
     repeat_region          4389. .4431
                            /rpt_family="(CA)n"
     repeat_region          5193. .5266
                            /rpt_family="L2"
     repeat_region          5556. .5645
                            /rpt_family="MIR"
     repeat_region          5999. .6169
                            /rpt_family="MER1_type"
     repeat_region          6202. .6316
                            /rpt_family="MIR"
     repeat_region          6555. .6606
                            /rpt_family="MIR"
     repeat_region          6910. .7272
                            /rpt_family="MaLR"
     repeat_region          7537. .7905
                            /rpt_family="MaLR"
     repeat_region          7909. .9510
                            /rpt_family="MaLR"
     repeat_region          9523. .9900
                            /rpt_family="MaLR"
     repeat_region          9925. .10132
                            /rpt_family="MaLR"
     repeat_region          10133. .10268
                            /rpt_family="Alu"
     repeat_region          10269. .10411
                            /rpt_family="MaLR"
     repeat_region          11024. .11061
                            /rpt_family="L2"
     repeat_region          11123. .11200
                            /rpt_family="MIR"
     repeat_region          11236. .11520
                            /rpt_family="L2"
     repeat_region          13157. .13347
                            /rpt_family="MIR"
     repeat_region          13360. .13772
                            /rpt_family="(ATGGTG)n"
     repeat_region          13754. .13929
                            /rpt_family="(TGG)n"
     repeat_region          13931. .13977
                            /rpt_family="(ATG)n"
     repeat_region          13987. .14222
                            /rpt_family="MIR"
     repeat_region          14423. .14527
                            /rpt_family="L2"
     repeat_region          14594. .14708
                            /rpt_family="L2"
     repeat_region          14855. .15035
                            /rpt_family="MIR"
     repeat_region          15426. .15727
                            /rpt_family="Alu"
     repeat_region          16021. .16126
                            /rpt_family="MIR"
     repeat_region          16429. .16856
                            /rpt_family="MaLR"
     repeat_region          17105. .17215
                            /rpt_family="MIR"
     repeat_region          17225. .17460
                            /rpt_family="L1"
     repeat_region          17461. .17820
```

```
                     /rpt_family="MaLR"
     repeat_region   17821. .19578
                     /rpt_family="L1"
     repeat_region   19582. .20133
                     /rpt_family="L1"
     repeat_region   20134. .20429
                     /rpt_family="Alu"
     repeat_region   20430. .21136
                     /rpt_family="L1"
     repeat_region   21282. .21533
                     /rpt_family="MIR"
     repeat_region   22088. .22402
                     /rpt_family="Alu"
     repeat_region   22757. .23239
                     /rpt_family="MaLR"
     repeat_region   23406. .23686
                     /rpt_family="ERVL"
     repeat_region   24004. .24189
                     /rpt_family="MaLR"
     repeat_region   24320. .24487
                     /rpt_family="MIR"
     repeat_region   24514. .24601
                     /rpt_family="L2"
     repeat_region   25075. .25182
                     /rpt_family="MIR"

  Query Match             1.7%;  Score 53;  DB 8;  Length 203583;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130670 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 130722


RESULT 36
CR847805
LOCUS       CR847805              204373 bp    DNA     linear   HTG 05-NOV-2004
DEFINITION  Homo sapiens chromosome 6 clone DAMA-118E22, 24 unordered pieces.
ACCESSION   CR847805
VERSION     CR847805.4  GI:55468491
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204373)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 5, 2004 this sequence version replaced gi:55294819.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bMA118E22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 195146 bases at least Q40
            Consensus quality: 198151 bases at least Q30
            Consensus quality: 199848 bases at least Q20
            Insert size: 202073; sum-of-contigs
            Insert size: 99271; 1.0% error; agarose-fp
            Quality coverage: 6.13x in Q20 bases; sum-of-contigs Quality
            coverage: 14.67x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 24 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3167: contig of 3167 bp in length
            *     3168     3267: gap of 100 bp
            *     3268     5570: contig of 2303 bp in length
            *     5571     5670: gap of 100 bp
            *     5671    13473: contig of 7803 bp in length
            *    13474    13573: gap of 100 bp
            *    13574    17619: contig of 4046 bp in length
            *    17620    17719: gap of 100 bp
            *    17720    20589: contig of 2870 bp in length
            *    20590    20689: gap of 100 bp
            *    20690    35318: contig of 14629 bp in length
            *    35319    35418: gap of 100 bp
            *    35419    40521: contig of 5103 bp in length
            *    40522    40621: gap of 100 bp
            *    40622    53837: contig of 13216 bp in length
            *    53838    53937: gap of 100 bp
            *    53938    56328: contig of 2391 bp in length
            *    56329    56428: gap of 100 bp
            *    56429    60179: contig of 3751 bp in length
            *    60180    60279: gap of 100 bp
            *    60280    63603: contig of 3324 bp in length
            *    63604    63703: gap of 100 bp
            *    63704    69568: contig of 5865 bp in length
            *    69569    69668: gap of 100 bp
            *    69669    72492: contig of 2824 bp in length
            *    72493    72592: gap of 100 bp
            *    72593    74831: contig of 2239 bp in length
            *    74832    74931: gap of 100 bp
            *    74932    77561: contig of 2630 bp in length
            *    77562    77661: gap of 100 bp
            *    77662    81419: contig of 3758 bp in length
            *    81420    81519: gap of 100 bp
            *    81520    86654: contig of 5135 bp in length
            *    86655    86754: gap of 100 bp
            *    86755    89390: contig of 2636 bp in length
            *    89391    89490: gap of 100 bp
            *    89491    94282: contig of 4792 bp in length
            *    94283    94382: gap of 100 bp
            *    94383   102792: contig of 8410 bp in length
            *   102793   102892: gap of 100 bp
            *   102893   112442: contig of 9550 bp in length
            *   112443   112542: gap of 100 bp
            *   112543   132047: contig of 19505 bp in length
            *   132048   132147: gap of 100 bp
            *   132148   162842: contig of 30695 bp in length
            *   162843   162942: gap of 100 bp
            *   162943   204373: contig of 41431 bp in length.
FEATURES             Location/Qualifiers
     source          1. .204373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMA-118E22"
                     /clone_lib="DNA-Arts BAC library MANN.1"
     misc_feature    1. .3167
                     /note="assembly_fragment:00230
                     fragment_chain:1"
     misc_feature    3268. .5570
                     /note="assembly_fragment:00124
                     fragment_chain:1"
     misc_feature    5671. .13473
                     /note="assembly_fragment:00481
                     fragment_chain:1"
     misc_feature    13574. .17619
                     /note="assembly_fragment:00293
```

```
                fragment_chain:1"
     misc_feature    17720. .20589
                     /note="assembly_fragment:00278
                     fragment_chain:2"
     misc_feature    20690. .35318
                     /note="assembly_fragment:00574
                     fragment_chain:2"
     misc_feature    35419. .40521
                     /note="assembly_fragment:00528
                     fragment_chain:2"
     misc_feature    40622. .53837
                     /note="assembly_fragment:00692
                     fragment_chain:2"
     misc_feature    53938. .56328
                     /note="assembly_fragment:00211
                     fragment_chain:3"
     misc_feature    56429. .60179
                     /note="assembly_fragment:00341
                     fragment_chain:3"
     misc_feature    60280. .63603
                     /note="assembly_fragment:00252
                     fragment_chain:4"
     misc_feature    63704. .69568
                     /note="assembly_fragment:00406
                     fragment_chain:4"
     misc_feature    69669. .72492
                     /note="assembly_fragment:00220"
     misc_feature    72593. .74831
                     /note="assembly_fragment:00241"
     misc_feature    74932. .77561
                     /note="assembly_fragment:00265"
     misc_feature    77662. .81419
                     /note="assembly_fragment:00308"
     misc_feature    81520. .86654
                     /note="assembly_fragment:00324"
     misc_feature    86755. .89390
                     /note="assembly_fragment:00358"
     misc_feature    89491. .94282
                     /note="assembly_fragment:00376"
     misc_feature    94383. .102792
                     /note="assembly_fragment:00437"
     misc_feature    102893. .112442
                     /note="assembly_fragment:00840.0"
     misc_feature    112543. .132047
                     /note="assembly_fragment:01041.0"
     misc_feature    132148. .162842
                     /note="assembly_fragment:01254.0"
     misc_feature    162943. .204373
                     /note="assembly_fragment:01615.0"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 204373;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156538 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 156590


RESULT 37
AC009107
LOCUS       AC009107               208220 bp    DNA     linear   PRI 14-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-459F6, complete sequence.
ACCESSION   AC009107
VERSION     AC009107.11  GI:28951164
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 208220)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 208220)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 208220)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 208220)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 14, 2003 this sequence version replaced gi:17298588.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .208220
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-459F6"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 8;  Length 208220;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14755 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 14807


RESULT 38
AC092434
LOCUS       AC092434               218872 bp    DNA     linear   HTG 27-JUL-2001
DEFINITION  Homo sapiens chromosome 4 clone RP11-210L21, *** SEQUENCING IN
            PROGRESS ***, 40 unordered pieces.
ACCESSION   AC092434
VERSION     AC092434.1  GI:14595956
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 218872)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 218872)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT
            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
```

```
Web site:http://genome.wustl.edu/gsc/index.shtml
-------------- Project Information --------------
Center project name: H_NH0210L21
-------------- Summary Statistics  --------------
Sequencing vector: M13; 14%
Sequencing vector: plasmid; 86%
Chemistry: Dye-primer ET; 14% of reads
Chemistry: Dye-terminator Big Dye; 86% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 188006 bases at least Q40
Consensus quality: 196601 bases at least Q30
Consensus quality: 202813 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 214972; sum-of-contigs
Quality coverage: 7.25 in Q20 bases; agarose-fp
Quality coverage: 5.64 in Q20 bases; sum-of-contigs
-------------------------------------------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     2669: contig of 2669 bp in length
*     2670     2769: gap of unknown length
*     2770     4554: contig of 1785 bp in length
*     4555     4654: gap of unknown length
*     4655     6908: contig of 2254 bp in length
*     6909     7008: gap of unknown length
*     7009    10387: contig of 3379 bp in length
*    10388    10487: gap of unknown length
*    10488    12825: contig of 2338 bp in length
*    12826    12925: gap of unknown length
*    12926    15663: contig of 2738 bp in length
*    15664    15763: gap of unknown length
*    15764    18992: contig of 3229 bp in length
*    18993    19092: gap of unknown length
*    19093    21884: contig of 2792 bp in length
*    21885    21984: gap of unknown length
*    21985    25307: contig of 3323 bp in length
*    25308    25407: gap of unknown length
*    25408    38126: contig of 12719 bp in length
*    38127    38226: gap of unknown length
*    38227    67816: contig of 29590 bp in length
*    67817    67916: gap of unknown length
*    67917   175443: contig of 107527 bp in length
*   175444   175543: gap of unknown length
*   175544   176878: contig of 1335 bp in length
*   176879   176978: gap of unknown length
*   176979   178260: contig of 1282 bp in length
*   178261   178360: gap of unknown length
*   178361   179616: contig of 1256 bp in length
*   179617   179716: gap of unknown length
*   179717   181151: contig of 1435 bp in length
*   181152   181251: gap of unknown length
*   181252   182971: contig of 1720 bp in length
*   182972   183071: gap of unknown length
*   183072   184476: contig of 1405 bp in length
*   184477   184576: gap of unknown length
*   184577   185838: contig of 1262 bp in length
*   185839   185938: gap of unknown length
*   185939   187304: contig of 1366 bp in length
*   187305   187404: gap of unknown length
*   187405   188505: contig of 1101 bp in length
*   188506   188605: gap of unknown length
*   188606   189895: contig of 1290 bp in length
*   189896   189995: gap of unknown length
*   189996   191186: contig of 1191 bp in length
*   191187   191286: gap of unknown length
*   191287   192476: contig of 1190 bp in length
*   192477   192576: gap of unknown length
*   192577   194011: contig of 1435 bp in length
*   194012   194111: gap of unknown length
*   194112   195415: contig of 1304 bp in length
*   195416   195515: gap of unknown length
*   195516   196928: contig of 1413 bp in length
*   196929   197028: gap of unknown length
*   197029   198215: contig of 1187 bp in length
*   198216   198315: gap of unknown length
*   198316   199494: contig of 1179 bp in length
*   199495   199594: gap of unknown length
*   199595   200810: contig of 1216 bp in length
*   200811   200910: gap of unknown length
*   200911   202155: contig of 1245 bp in length
*   202156   202255: gap of unknown length
*   202256   204047: contig of 1792 bp in length
*   204048   204147: gap of unknown length
*   204148   205275: contig of 1128 bp in length
*   205276   205375: gap of unknown length
*   205376   207824: contig of 2449 bp in length
*   207825   207924: gap of unknown length
*   207925   209566: contig of 1642 bp in length
*   209567   209666: gap of unknown length
*   209667   211262: contig of 1596 bp in length
*   211263   211362: gap of unknown length
*   211363   213638: contig of 2276 bp in length
*   213639   213738: gap of unknown length
*   213739   215671: contig of 1933 bp in length
*   215672   215771: gap of unknown length
*   215772   216917: contig of 1146 bp in length
*   216918   217017: gap of unknown length
*   217018   218872: contig of 1855 bp in length.
FEATURES             Location/Qualifiers
     source          1. .218872
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="RP11-210L21"
     misc_feature    1. .2669
                     /note="assembly_name:Contig100"
     gap             2670. .2769
                     /estimated_length=unknown
     misc_feature    2770. .4554
                     /note="assembly_name:Contig101"
     gap             4555. .4654
                     /estimated_length=unknown
     misc_feature    4655. .6908
                     /note="assembly_name:Contig102"
     gap             6909. .7008
                     /estimated_length=unknown
     misc_feature    7009. .10387
                     /note="assembly_name:Contig103"
     gap             10388. .10487
                     /estimated_length=unknown
     misc_feature    10488. .12825
                     /note="assembly_name:Contig104"
     gap             12826. .12925
                     /estimated_length=unknown
     misc_feature    12926. .15663
                     /note="assembly_name:Contig105"
     gap             15664. .15763
                     /estimated_length=unknown
     misc_feature    15764. .18992
                     /note="assembly_name:Contig106"
     gap             18993. .19092
                     /estimated_length=unknown
     misc_feature    19093. .21884
                     /note="assembly_name:Contig107"
     gap             21885. .21984
                     /estimated_length=unknown
     misc_feature    21985. .25307
                     /note="assembly_name:Contig108"
     gap             25308. .25407
```

```
                     /estimated_length=unknown
     misc_feature    25408. .38126
                     /note="assembly_name:Contig109"
     gap             38127. .38226
                     /estimated_length=unknown
     misc_feature    38227. .67816
                     /note="assembly_name:Contig110
                     clone_end:T7
                     vector_side:left"
     gap             67817. .67916
                     /estimated_length=unknown
     misc_feature    67917. .175443
                     /note="assembly_name:Contig111
                     clone_end:SP6
                     vector_side:right"
     gap             175444. .175543
                     /estimated_length=unknown
     misc_feature    175544. .176878
                     /note="assembly_name:Contig42"
     gap             176879. .176978
                     /estimated_length=unknown
     misc_feature    176979. .178260
                     /note="assembly_name:Contig46"
     gap             178261. .178360
                     /estimated_length=unknown
     misc_feature    178361. .179616
                     /note="assembly_name:Contig51"
     gap             179617. .179716
                     /estimated_length=unknown

  Query Match             1.7%;  Score 53;  DB 14;  Length 218872;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195757 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 195809


RESULT 39
CR847784
LOCUS       CR847784               229090 bp    DNA     linear   HTG 05-NOV-2004
DEFINITION  Homo sapiens chromosome 6 clone DAMA-182N14, 10 unordered pieces.
ACCESSION   CR847784
VERSION     CR847784.4  GI:55468488
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 229090)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 5, 2004 this sequence version replaced gi:54035014.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bMA182N14
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 225613 bases at least Q40
            Consensus quality: 226328 bases at least Q30
            Consensus quality: 226997 bases at least Q20
            Insert size: 228190; sum-of-contigs
            Insert size: 111318; 0.8% error; agarose-fp
            Quality coverage: 8.67x in Q20 bases; sum-of-contigs Quality
            coverage: 19.14x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *      1    24138: contig of 24138 bp in length
            *  24139    24238: gap of 100 bp
            *  24239    47128: contig of 22890 bp in length
            *  47129    47228: gap of 100 bp
            *  47229    66763: contig of 19535 bp in length
            *  66764    66863: gap of 100 bp
            *  66864    72248: contig of 5385 bp in length
            *  72249    72348: gap of 100 bp
            *  72349    77388: contig of 5040 bp in length
            *  77389    77488: gap of 100 bp
            *  77489    88555: contig of 11067 bp in length
            *  88556    88655: gap of 100 bp
            *  88656    98102: contig of 9447 bp in length
            *  98103    98202: gap of 100 bp
            *  98203   129421: contig of 31219 bp in length
            * 129422   129521: gap of 100 bp
            * 129522   187227: contig of 57706 bp in length
            * 187228   187327: gap of 100 bp
            * 187328   229090: contig of 41763 bp in length.
FEATURES             Location/Qualifiers
     source          1. .229090
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMA-182N14"
                     /clone_lib="DNA-Arts BAC library MANN.1"
     misc_feature    1. .24138
                     /note="assembly_fragment:00874
                     fragment_chain:1"
     misc_feature    24239. .47128
                     /note="assembly_fragment:00462
                     fragment_chain:1"
     misc_feature    47229. .66763
                     /note="assembly_fragment:00286
                     fragment_chain:1"
     misc_feature    66864. .72248
                     /note="assembly_fragment:00093"
     misc_feature    72349. .77388
                     /note="assembly_fragment:00121"
     misc_feature    77489. .88555
                     /note="assembly_fragment:00192.0"
     misc_feature    88656. .98102
                     /note="assembly_fragment:00657.0"
     misc_feature    98203. .129421
                     /note="assembly_fragment:01123.0"
     misc_feature    129522. .187227
                     /note="assembly_fragment:01753.0"
     misc_feature    187328. .229090
                     /note="assembly_fragment:02711.0"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 229090;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123117 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 123169


RESULT 40
```

```
CR753824
LOCUS       CR753824               323185 bp    DNA     linear   HTG 05-NOV-2004
DEFINITION  Homo sapiens chromosome 6 clone DADB-56O24, *** SEQUENCING IN
            PROGRESS ***, 4 unordered pieces.
ACCESSION   CR753824
VERSION     CR753824.2  GI:55468370
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 323185)
  AUTHORS   Leongamornlert,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 5, 2004 this sequence version replaced gi:51571640.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bDB56O24
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 89730 bases at least Q40
            Consensus quality: 89774 bases at least Q30
            Consensus quality: 89790 bases at least Q20
            Insert size: 322885; sum-of-contigs
            Insert size: 96449; 1.4% error; agarose-fp
            Quality coverage: 1.82x in Q20 bases; sum-of-contigs Quality
            coverage: 6.22x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    15449: contig of 15449 bp in length
            *   15450    15549: gap of 100 bp
            *   15550    60747: contig of 45198 bp in length
            *   60748    60847: gap of 100 bp
            *   60848   165052: contig of 104205 bp in length
            *  165053   165152: gap of 100 bp
            *  165153   323185: contig of 158033 bp in length.
FEATURES             Location/Qualifiers
     source          1. .323185
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DADB-56O24"
                     /clone_lib="DNA-arts-BAC.1-DBB.1"
     misc_feature    1. .15449
                     /note="assembly_fragment:00142
                     fragment_chain:1"
     misc_feature    15550. .60747
                     /note="assembly_fragment:00419
                     fragment_chain:1"
     misc_feature    60848. .165052
                     /note="assembly_fragment:00920"
     misc_feature    165153. .323185
                     /note="assembly_fragment:00949"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 14;  Length 323185;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288592 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 288644


RESULT 41
AP001710
LOCUS       AP001710               340000 bp    DNA     linear   PRI 21-MAY-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 21q, section 54/105.
ACCESSION   AP001710 AL163255 BA000005
VERSION     AP001710.1  GI:7768781
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
  TITLE     The DNA sequence of human chromosome 21
  JOURNAL   Nature 405 (6784), 311-319 (2000)
   PUBMED   10830953
REFERENCE   2  (bases 1 to 340000)
  AUTHORS   Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K.,
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing
            Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
            Group * Institute of Molecular Biotechnology, Genome Analysis *
            Keio University School of Medicine, Dept. of Molecular Biology *
            GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
            Genetics (addresses see below)
COMMENT     On May 30, 2000 this sequence version replaced gi:7717322.
            The chromosome 21 mapping and sequencing consortium consisting of
            * RIKEN Genomic Sciences Center, Human Genome Research Group, *
            Sagamihara 228-8555, Japan,
            * e.mail: hattori@gsc.riken.go.jp
            * URL: http://hgp.gsc.riken.go.jp/
            and
            * Institute of Molecular Biotechnology, Genome Analysis, *
            Beutenbergstrasse 11, D-07745 Jena, Germany,
            * e.mail: gscj-submit@genome.imb-jena.de
            * URL: http://genome.imb-jena.de/
            and
            * Keio University School of Medicine, Molecular Biology, * Tokyo
            160-8582, Japan,
            * e.mail: nshimizu@dmb-med.keio.ac.jp
```

```
          * URL: http://www.dmb.med.keio.ac.jp/
          and
          * GBF, Dept. of Genome Analysis,
          * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
          info.genome@gbf.de
          * URL: http://genome.gbf.de/
          and
          * Max-Planck Institute for Molecular Genetics,
          * Ihnestrasse 73, D-14195 Berlin, Germany,
          * e.mail: info-chr21@molgen.mpg.de
          * URL: http://chr21.rz-berlin.mpg.de/
          AL163255: Submitted (10-Apr-2000).
FEATURES             Location/Qualifiers
     source          1. .340000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     source          <1. .52327
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R14B21, 5' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000245"
     source          51884. .146095
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R137B7"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000246"
     source          145455. .259373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="PAC871G18"
                     /clone_lib="RP5 PAC library"
                     /note="Accession No. AP000247"
     source          256917. .>340000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1"
                     /clone="R289H18, 3' partial"
                     /clone_lib="RPCI-11 BAC library"
                     /note="Accession No. AP000248"
     repeat_region   complement(783. .1089)
                     /note="AluSc"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(3359. .3661)
                     /note="AluJo"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   4287. .4321
                     /note="(T)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   5078. .5654
                     /note="MER74A"
                     /rpt_family="LTR/MER73-group"
                     /rpt_type=DISPERSED
     repeat_region   6855. .7314
                     /note="LTR2"
                     /rpt_family="LTR/Retroviral"
                     /rpt_type=DISPERSED
     repeat_region   7666. .7725
                     /note="(TA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   7761. .7825
                     /note="(TA)n"
                     /rpt_family="Simple_repeat"
                     /rpt_type=TANDEM
     repeat_region   8223. .8411
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   8625. .9032
                     /note="L1M4"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   9504. .10065
                     /note="L1ME3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   complement(10066. .10198)
                     /note="FLAM_C"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   10199. .10664
                     /note="L1ME3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   10690. .10950
                     /note="L1ME3"
                     /rpt_family="LINE/L1"
                     /rpt_type=DISPERSED
     repeat_region   11012. .11310
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   11596. .11903
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(13161. .13475)
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(13777. .14219)
                     /note="MLT1E2"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   14220. .14505
                     /note="AluY"
                     /rpt_family="SINE/Alu"
                     /rpt_type=DISPERSED
     repeat_region   complement(14506. .14554)
                     /note="MLT1E2"
                     /rpt_family="LTR/MaLR"
                     /rpt_type=DISPERSED
     repeat_region   14670. .14795
                     /note="MIR"
                     /rpt_family="SINE/MIR"
                     /rpt_type=DISPERSED
     repeat_region   14838. .14909
                     /note="L2"
                     /rpt_family="LINE/L2"

  Query Match             1.7%;  Score 53;  DB 8;  Length 340000;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     159396 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 159448
```

```
RESULT 42
CS039425
LOCUS       CS039425              349980 bp    DNA     linear   PAT 22-MAR-2005
DEFINITION  Sequence 235 from Patent WO2005019477.
ACCESSION   CS039425
VERSION     CS039425.1  GI:61847452
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lewin,J., Berlin,K., Hildmann,T., Olek,A., Beck,S. and Novik,K.
  TITLE     Methods and compositions for differentiating tissues or cell types
            using epigenetic markers
  JOURNAL   Patent: WO 2005019477-A 235 03-MAR-2005;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="unknown base
                     seq 205 to long, 4.267.840
                     replaced by
                     new seq 205, from 0.000.001 to 0.349.980
                     new seq 222, from 0.300.001 to 0.649.980
                     new seq 223, from 0.600.001 to 0.949.980
                     new seq 224, from 0.900.001 to 1.249.980
                     new seq 225, from 1.200.001 to 1.549.980
                     new seq 226, from 1.500.001 to 1.849.980
                     new seq 227, from 1.800.001 to 2.149.980
                     new seq 228, from 2.100.001 to 2.449.980
                     new seq 229, from 2.400.001 to 2.749.980
                     new seq 230, from 2.700.001 to 3.049.980
                     new seq 231, from 3.000.001 to 3.349.980
                     new seq 232, from 3.300.001 to 3.649.980
                     new seq 233, from 3.600.001 to 3.949.980
                     new seq 234, from 3.900.001 to 4.249.980
                     new seq 235, from 4.200.001 to 4.549.980
                     new seq 236, from 4.500.001 to 4.647.455"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 6;  Length 349980;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240219 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 240271


RESULT 43
BD027540/c
LOCUS       BD027540                381 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD027540
VERSION     BD027540.1  GI:22569282
KEYWORDS    JP 2001269182-A/3786.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 3786 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/3786
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC   score 7.7
            CC   seq LCLLGSSDSPASA/SQ
            FH   Key             Location/Qualifiers
            FT   CDS             12. .347
            FT   sig_peptide     12. .158
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 56


RESULT 44
AX887930/c
LOCUS       AX887930                381 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 3793 from Patent EP1033401.
ACCESSION   AX887930
VERSION     AX887930.1  GI:40046583
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 3793 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             12. .350
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE93855.1"
                     /db_xref="GI:40046584"
                     /translation="MHVVVVVVVVVVVVVLRQSLALLPRLECSGTILAHCSLCLLGSSD
                     SPASASQVARTTGMCHHTQLIFVFLIETGFLHIGQAALELLTSGDPPASASQSAGITG
                     VXTTPGRLSN"
     sig_peptide     12. .158
                     /note="score 7.7 seq LCLLGSSDSPASA/SQ"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 56
```

```
RESULT 45
AB135735/c
LOCUS       AB135735                428 bp    DNA     linear   STS 02-APR-2005
DEFINITION  Homo sapiens DNA, STS on chromosome 1, D1S0306i, sequence tagged
            site.
ACCESSION   AB135735
VERSION     AB135735.1  GI:62156316
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tamiya,G., Shinya,M., Ikuta,T., Makino,S., Matsumoto,T., Mano,S.,
            Furugaki,K., Ando,S., Nozaki,Y., Yukawa,W., Nakashige,R.,
            Yamaguchi,D., Ishibashi,H., Yonekura,M., Nakami,Y., Takayama,S.,
            Endo,T., Saruwatari,T., Yagura,M., Yoshikawa,Y., Fujimoto,K.,
            Oka,A., Chiku,S., Linsen,S.E., Giphart,M.J., Bahram,S.,
            Kulski,Y.J., Fukazawa,T., Hashimoto,H., Hoshina,Y., Suzuki,Y.,
            Hotta,T., Mochida,J., Minezaki,T., Komai,K., Shiozawa,S.,
            Taniguchi,A., Yamanaka,H., Kamatani,N., Imanishi,T., Gojobori,T.
            and Inoko,H.
  TITLE     A Whole Genome Association Study of Rheumatoid Arthritis using
            27,039 Microsatellite Makers
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 428)
  AUTHORS   Tamiya,G., Makino,S., Fujimoto,K., Oka,A., Hayashi,H., Denda,A.,
            Linsen,S.E., Ikuta,T., Shinya,M., Endo,T., Tomizawa,M., Tokubo,E.,
            Sato,R., Takaki,A., Nagatsuka,Y., Watanabe,H., Adachi,S.,
            Makino,Y., Nakano,S., Yamamoto,A., Yoshida,K., Okamoto,K.,
            Yamaguchi,D., Ishibashi,H., Yonekura,M., Takayama,S., Nakami,Y.,
            Saruwatari,T., Brand,A., van Hilten,J.A., van de Watering,L.M.,
            Giphart,M.J., Bahram,S., Kulski,Y.J. and Inoko,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-DEC-2003) Hidetoshi Inoko, Tokai University School of
            Medicine, Department of Genetic Information; Bohseidai, Isehara,
            Kanagawa, 259-1193, Japan (E-mail:hinoko@is.icc.u-tokai.ac.jp,
            Tel:81-463-93-1121, Fax:81-463-94-8884)
COMMENT     Polymorphisms were confirmed by comparing with the pooled DNA
            typing data of 88 Dutch population or 100 Australian. When there
            was no significant difference between the Japanese and these
            Caucasians, individual typing was performed to confirm
            polymorphisms.
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1p36.23"
                     /note="pooled DNA of 100 Japanese unrelated individuals
                     sequence tagged site D1S0306i"
     primer_bind     1. .24
                     /note="5' primer: AACTGTAGAATCTAAACGGTGAGT"
                     /PCR_conditions="denaturation 96degC 5 min, 57degC 1 min,
                     72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec,
                     72degC 1 min"
     repeat_region   80. .111
                     /note="sequence tagged site D1S0306i"
                     /rpt_type=tandem
                     /rpt_unit="gttt"
     primer_bind     complement(405. .428)
                     /note="3' primer: AAACATGATGCTCACTTTAAGAAC"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 116
```

```
RESULT 46
BV640976/c
LOCUS       BV640976                690 bp    DNA     linear   STS 16-APR-2005
DEFINITION  S217P60133RD8.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV640976
VERSION     BV640976.1  GI:62668946
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 690
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>690
```

ORIGIN

Query Match 1.7%; Score 52; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 5.1e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 212 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 161

RESULT 47
BV533342
LOCUS BV533342 762 bp DNA linear STS 08-APR-2005
DEFINITION G591P626298RE9.T0 Clint Pan troglodytes verus STS genomic, sequence tagged site.
ACCESSION BV533342
VERSION BV533342.1 GI:62411126
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 762)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 762
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.
FEATURES Location/Qualifiers
source 1. .762
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>762
ORIGIN

Query Match 1.7%; Score 52; DB 10; Length 762;
Best Local Similarity 100.0%; Pred. No. 5.1e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 59

RESULT 48
BV536183
LOCUS BV536183 770 bp DNA linear STS 08-APR-2005
DEFINITION G591P631495FD4.T0 Clint Pan troglodytes verus STS genomic, sequence tagged site.
ACCESSION BV536183
VERSION BV536183.1 GI:62413977
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 770)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 770
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied

```
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES              Location/Qualifiers
     source           1. .770
                      /organism="Pan troglodytes verus"
                      /mol_type="genomic DNA"
                      /sub_species="verus"
                      /db_xref="taxon:37012"
                      /clone_lib="Clint"
     STS              <1. .>770
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 72


RESULT 49
BV588891
LOCUS       BV588891                 773 bp    DNA     linear   STS 12-APR-2005
DEFINITION  G591P638732FA4.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV588891
VERSION     BV588891.1  GI:62504620
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 773
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES              Location/Qualifiers
     source           1. .773
                      /organism="Pan troglodytes verus"
                      /mol_type="genomic DNA"
                      /sub_species="verus"
                      /db_xref="taxon:37012"
                      /clone_lib="Clint"
     STS              <1. .>773
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 773;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 84


RESULT 50
BV492390
LOCUS       BV492390                 775 bp    DNA     linear   STS 07-APR-2005
DEFINITION  S221P60653FF7.T0 Yvonne Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV492390
VERSION     BV492390.1  GI:62336060
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 775
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
```

```
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Yvonne"
     STS             <1. .>775
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 775;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 459


RESULT 51
BV523612/c
LOCUS       BV523612                 810 bp    DNA     linear   STS 08-APR-2005
DEFINITION  G591P600782FH3.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV523612
VERSION     BV523612.1  GI:62401382
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 810
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>810
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 810;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 341


RESULT 52
BV465101
LOCUS       BV465101                 915 bp    DNA     linear   STS 06-APR-2005
DEFINITION  G591P61938RC7.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV465101
VERSION     BV465101.1  GI:62224529
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
```

```
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 915
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>915
ORIGIN

  Query Match             1.7%;  Score 52;  DB 10;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 702


RESULT 53
AB014078/c
LOCUS       AB014078               22738 bp    DNA     linear   PRI 20-NOV-1999
DEFINITION  Homo sapiens genomic DNA, chromosome 6p21.3, HLA class I region,
            Cosmid clone:TY1C6, complete sequence.
ACCESSION   AB014078
VERSION     AB014078.1  GI:5672590
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Shiina,T., Tamiya,G., Oka,A., Takishima,N., Yamagata,T.,
            Kikkawa,E., Iwata,K., Tomizawa,M., Okuaki,N., Kuwano,Y.,
            Watanabe,K., Fukuzumi,Y., Itakura,S., Sugawara,C., Ono,A.,
            Yamazaki,M., Tashiro,H., Ando,A., Ikemura,T., Soeda,E., Kimura,M.,
            Bahram,S. and Inoko,H.
  TITLE     Molecular dynamics of MHC genesis unraveled by sequence analysis of
            the 1,796,938-bp HLA class I region
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (23), 13282-13287 (1999)
   PUBMED   10557312
REFERENCE   2  (bases 1 to 22738)
  AUTHORS   Shiina,T., Tamiya,G., Oka,A. and Takishima,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-1998) Nobusada Takishima, Tokai University School
            of Medicine, Molecular Life Science; Bohseidai, Isehara, Kanagawa
            259-11, Japan (E-mail:takisan@is.icc.u-tokai.ac.jp,
            Tel:81-463-93-1121, Fax:81-463-94-8884)
FEATURES             Location/Qualifiers
     source          1. .22738
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
                     /clone="745D12 and 960H11"
                     /haplotype="HLA-DR4"
                     /sex="male"
                     /cell_line="Boleth"
                     /cell_type="Immunoresponse cell"
                     /tissue_type="blood"
                     /clone_lib="CEPH YAC"
                     /dev_stage="adult"
                     /note="Cosmid Clone TY1C6"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 22738;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19661 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 19610


RESULT 54
BX248096
LOCUS       BX248096               23481 bp    DNA     linear   PRI 28-MAY-2004
DEFINITION  Human DNA sequence from clone DASS-205E5 on chromosome 6, complete
            sequence.
ACCESSION   BX248096
VERSION     BX248096.10  GI:47824849
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 23481)
  AUTHORS   Wood,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 28, 2004 this sequence version replaced gi:47776042.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
```

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6/MHC
DASS-205E5 is from a DNA-arts SSTO human bac library VECTOR: pBeloBAC11.

FEATURES Location/Qualifiers
source 1. .23481
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DASS-205E5"
/clone_lib="DNA-arts-BAC.1-SSTO.1"

ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 23481;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7576 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7627

RESULT 55
AC087649/c
LOCUS AC087649 25057 bp DNA linear PRI 07-MAY-2002
DEFINITION Homo sapiens chromosome 17, clone RP11-46I22, complete sequence.
ACCESSION AC087649
VERSION AC087649.6 GI:20486426
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 25057)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-46I22
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 25057)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 25057)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 25057)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On May 7, 2002 this sequence version replaced gi:20455562.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

```
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
          ------------- Project Information
            Center project name: L12000
            Center clone name: 46_I_22
          -------------

          Only the middle 25.1 kilobases of this clone are being submitted.
          The remainder overlaps either accession number AC068234 [WICGR
          project L10228]
          or accession number AC002558 [WICGR project L239].
FEATURES             Location/Qualifiers
     source          1. .25057
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-46I22"
                     /clone_lib="RPCI-11 Human Male BAC"
     unsure          1. .11
                     /note="<30 qual single clone coverage"
     repeat_region   2. .70
                     /rpt_family="MIR"
     unsure          13. .22
                     /note="<30 qual SNGL region"
     unsure          51. .56
                     /note="<30 qual SNGL region"
     repeat_region   185. .495
                     /rpt_family="AluSp"
     repeat_region   complement(548. .665)
                     /rpt_family="MIR"
     repeat_region   864. .1097
                     /rpt_family="L3"
     repeat_region   1098. .1404
                     /rpt_family="AluJo"
     repeat_region   1405. .1498
                     /rpt_family="L3"
     repeat_region   complement(1521. .1648)
                     /rpt_family="MIR"
     repeat_region   1689. .1853
                     /rpt_family="L3"
     repeat_region   2191. .2522
                     /rpt_family="L1MC4a"
     repeat_region   2628. .2689
                     /rpt_family="MIR3"
     repeat_region   2980. .3236
                     /rpt_family="MIR"
     repeat_region   3272. .3305
                     /rpt_family="(TCC)n"
     repeat_region   complement(3582. .3873)
                     /rpt_family="MER2"
     repeat_region   4236. .4646
                     /rpt_family="LTR54"
     repeat_region   4754. .4840
                     /rpt_family="L2"
     repeat_region   complement(4932. .5234)
                     /rpt_family="AluSq"
     repeat_region   5286. .5316
                     /rpt_family="(CATCC)n"
     repeat_region   complement(5774. .5911)
                     /rpt_family="AluJo"
     repeat_region   7663. .7712
                     /rpt_family="(CTGGGG)n"
     repeat_region   8499. .8584
                     /rpt_family="MIR3"
     repeat_region   8588. .8626
                     /rpt_family="(TG)n"
     repeat_region   complement(8627. .8718)
                     /rpt_family="MIR"
     repeat_region   complement(8719. .9044)
                     /rpt_family="AluJo"
     repeat_region   complement(9045. .9185)
                     /rpt_family="MIR"
     repeat_region   9546. .9644
                     /rpt_family="MIR3"
     repeat_region   complement(9718. .9876)
                     /rpt_family="MIR3"
     repeat_region   10345. .10450
                     /rpt_family="MIR"
     repeat_region   10976. .11028
                     /rpt_family="MIR"
     repeat_region   12364. .12427
                     /rpt_family="MIR"
     repeat_region   complement(12430. .12599)
                     /rpt_family="AluJb"
     repeat_region   12600. .12731
                     /rpt_family="MIR"
     repeat_region   13535. .13922
                     /rpt_family="L1M4"
     repeat_region   13951. .14236
                     /rpt_family="AluSx"
     repeat_region   14239. .14891
                     /rpt_family="Charlie5"
     repeat_region   14892. .15194
                     /rpt_family="AluSx"
     repeat_region   15195. .15269
                     /rpt_family="Charlie5"
     repeat_region   15422. .15731
                     /rpt_family="AluSp"
     repeat_region   complement(15984. .16287)
                     /rpt_family="AluSx"

  Query Match             1.7%;  Score 52;  DB 8;  Length 25057;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17988 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 17937


RESULT 56
AL356788
LOCUS       AL356788               25242 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-308L13 on chromosome 1, complete
            sequence.
ACCESSION   AL356788 AC068484
VERSION     AL356788.6  GI:15626410
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 25242)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 16, 2001 this sequence version replaced gi:13398766.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-308L13 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            ------------- Genome Center
```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vega@sanger.ac.uk

Draft Sequence Produced by Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
http://genome.wustl.edu/gsc/index.shtml
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```
FEATURES             Location/Qualifiers
     source          1. .25242
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-308L13"
                     /clone_lib="RPCI-11.2"
     misc_feature    2000
                     /note="Clone_right_end: RP11-103C3"
     misc_feature    23243
                     /note="Clone_left_end: RP11-191N8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 25242;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3472 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3523


RESULT 57
CR936916
LOCUS       CR936916               25972 bp    DNA     linear   PRI 03-MAR-2005
DEFINITION  Human DNA sequence from clone DAMC-259M15 on chromosome 6, complete
            sequence.
ACCESSION   CR936916
VERSION     CR936916.2  GI:60495381
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 25972)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 3, 2005 this sequence version replaced gi:60457977.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ---
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAMC-259M15 is from the DNA-Arts.org BAC library MCF.1 VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .25972
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMC-259M15"
                     /clone_lib="DNA-Arts.org BAC library MCF.1"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 25972;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24111 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 24162


RESULT 58
CR382280
LOCUS       CR382280               28276 bp    DNA     linear   PRI 08-JUN-2004
DEFINITION  Human DNA sequence from clone DAMA-277I14 on chromosome 6, complete
            sequence.
ACCESSION   CR382280
VERSION     CR382280.5  GI:48473864
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 28276)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 8, 2004 this sequence version replaced gi:47604309.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ---
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
```

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6/MHC
DAMA-277I14 is from the DNA-Arts human BAC library MANN.1 VECTOR: pBeloBAC11.

```
FEATURES             Location/Qualifiers
     source          1. .28276
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMA-277I14"
                     /clone_lib="DNA-Arts BAC library MANN.1"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 28276;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26981 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 27032


RESULT 59
CR391992/c
LOCUS       CR391992               32906 bp    DNA     linear   PRI 19-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-308E4 on chromosome 1 Contains a
            novel gene (FLJ13171) and a DnaJ (Hsp40) homolog subfamily C member
            8 (DNAJC8) pseudogene, complete sequence.
ACCESSION   CR391992
VERSION     CR391992.5  GI:47232524
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 32906)
  AUTHORS   Barlow,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 14, 2004 this sequence version replaced gi:47058867.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP1-308E4 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .32906
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP1-308E4"
                     /clone_lib="RPCI-1"
     misc_feature    1
                     /note="Clone_left_end: RP1-308E4"
     gene            3893. .4306
                     /locus_tag="RP1-308E4.1-001"
                     /pseudo
     CDS             3893. .4306
                     /locus_tag="RP1-308E4.1-001"
                     /note="match: proteins: AAH66222 Sw:O75937 Tr:AAH57191
                     Tr:Q7QI49 Tr:Q8C2M6 Tr:Q8N4Z5 Tr:Q9VCH9"
                     /pseudo
                     /codon_start=1
                     /product="DnaJ (Hsp40) homolog, subfamily C member 8
                     (DNAJC8) pseudogene"
     gene            join(8205. .8232,9458. .9631,10585. .10665,16058. .16222,
                     16755. .17141,23928. .24525,26481. .26665,30825. .30978,
                     AL840643.1:8638. .8693,AL840643.1:9242. .9362,
                     AL840643.1:10419. .10480,AL840643.1:11118. .11255,
                     AL840643.1:12483. .12559,AL840643.1:13698. .13790,
                     AL840643.1:16002. .16624)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-002"
     mRNA            join(8205. .8232,9458. .9631,10585. .10665,16058. .16222,
                     16755. .17141,23928. .24525,26481. .26665,30825. .30978,
                     AL840643.1:8638. .8693,AL840643.1:9242. .9362,
                     AL840643.1:10419. .10480,AL840643.1:11118. .11255,
                     AL840643.1:12483. .12559,AL840643.1:13698. .13790,
                     AL840643.1:16002. .16624)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-002"
                     /product="novel protein"
                     /note="match: ESTs: Em:BG722393.1 Em:CN346119.1
                     match: cDNAs: Em:BC068508.1"
     gene            join(9458. .9631,10585. .10665,16058. .16222,16755. .17141,
                     23928. .24525,26481. .26665,30825. .30978,
                     AL840643.1:8638. .8693,AL840643.1:10419. .10480,
                     AL840643.1:11118. .11255,AL840643.1:12483. .12559,
                     AL840643.1:16002. .19837)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-001"
     mRNA            join(9458. .9631,10585. .10665,16058. .16222,16755. .17141,
                     23928. .24525,26481. .26665,30825. .30978,
                     AL840643.1:8638. .8693,AL840643.1:10419. .10480,
                     AL840643.1:11118. .11255,AL840643.1:12483. .12559,
                     AL840643.1:16002. .19837)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-001"
                     /product="novel protein"
                     /note="match: ESTs: Em:AU132111.1 Em:AU134282.1
                     Em:BF794049.1 Em:BG170714.1
                     match: cDNAs: Em:AF130081.1 Em:AK023233.1 Em:AK025436.1
                     Em:BC021247.2 Em:BC029266.1"
     gene            join(9458. .9631,10585. .10665,16058. .16222,16755. .17614)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-003"
     mRNA            join(9458. .9631,10585. .10665,16058. .16222,16755. .17614)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-003"
                     /product="novel protein"
```

```
                     /note="match: cDNAs: Em:BC016152.1"
     CDS             join(9490. .9631,10585. .10665,16058. .16222,16755. .17141,
                     23928. .24525,26481. .26665,30825. .30978,
                     AL840643.1:8638. .8693,AL840643.1:10419. .10480,
                     AL840643.1:11118. .11255,AL840643.1:12483. .12559,
                     AL840643.1:16002. .16017)
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-001"
                     /standard_name="OTTHUMP00000003648"
                     /note="match: proteins: Sw:O75167 Tr:Q8IZ21 Tr:Q8N384
                     Tr:Q91VK6 Tr:Q9H395 Tr:Q9H6X0 Tr:Q9H8W6"
                     /codon_start=1
                     /product="novel protein"
                     /protein_id="CAI20111.1"
                     /db_xref="GI:56208555"
                     /db_xref="InterPro:IPR004018"
                     /db_xref="UniProt/TrEMBL:Q8N384"
                     /translation="MVLDSVEAGDTTPPTKRKSKFSGFGKIFKPWKWRKKKSSDKFKE
                     TSEVLERKISMRKPREELVKRGVLLEDPEQGGEDPGKPSDAMLKNGHTTPIGNARSSS
                     PVQVEEEPVRLASLRKAIPEEDLKKRLGSTGSQPNSEAESVPENVPKPPLLPPKRPLS
                     SSHEASEGQAKDATSSGGTARFIISTSITTAPAATTAATSLAKTVNLSVTPSPAPRTL
                     PAAPASTNTTATPSLTHMVPAKQPPIPPPKPAHRNSNPVIAELSQAINSGTLLSKPSP
                     PLPPKRGIPSTSVPTLESAAAITTKTPSDEREKSTCSMGSELLPMISPRSPSPPLPTH
                     IPPEPPRTPPPFAKTFQVVPEIEFPPSLDLHQEIPQQEDQKKEVPKRILDQNFGEPHI
                     PSRLPPLPLHIRIQQALTSPLPMTPILEGSHRAHSLLFENSDSFSEDSSTLGRTRSLP
                     ITIEMLKVPDDEEEEEQTCPSTFSEEMTPTSVIPKLPQCLREEEEKESDSDSEGPIQY
                     RDEEDEDESYQSALANKVKRKDTLAMKLNHRPSEPELNLNSWPCKSKEEWNEIRHQIG
                     NTLIRRLSQRPTPEELEQRNILQPKNEADRQAEKREIKRRLTRKLSQRPTVAELLARK
                     ILRFNEYVEVTDAQDYDRRADKPWTKLTPADKAAIRKELNEFKSSEMEVHEESKHFTR
                     YHRP"
     misc_feature    30907
                     /gene="RP11-442N24__A.1"
                     /locus_tag="RP11-442N24__A.1-001"
                     /note="Clone_left_end: RP11-442N24"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 32906;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10951 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 10900


RESULT 60
AC093235/c
LOCUS       AC093235              38173 bp    DNA     linear   PRI 29-SEP-2001
DEFINITION  Homo sapiens chromosome 19 clone LLNLR-267A4, complete sequence.
ACCESSION   AC093235
VERSION     AC093235.2  GI:15808549
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 29, 2001 this sequence version replaced gi:15193369.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .38173
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="LLNLR-267A4"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 38173;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19282 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 19231


RESULT 61
HSN121E8
LOCUS       HSN121E8              39372 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone LL22NC03-121E8 on chromosome 22,
            complete sequence.
ACCESSION   Z78421
VERSION     Z78421.1  GI:1495467
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 39372)
  AUTHORS   Kettleborough,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep LL22NC03-121E8
            is from the human chromosome 22-specific cosmid library LL22NC03,
            constructed at the Biomedical Sciences Division, Lawrence Livermore
            National Laboratory, Livermore, CA 94550 under the auspices of the
            National Laboratory Gene Library Project sponsored by the US
            Department of Energy.  The source of the flow sorted chromosomes
            was a human/hamster hybrid containing chromosomes Y, 22 and 9.
            VECTOR: lawrist16
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .39372
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /clone="LL22NC03-121E8"
                     /clone_lib="LL22NC03"
     gene            complement(join(5697. .5931,6975. .7281))
                     /locus_tag="LL22NC03-121E8.2-002"
     mRNA            complement(join(5697. .5931,6975. .7281))
                     /locus_tag="LL22NC03-121E8.2-002"
     gene            join(7219. .8044,8241. .8279,8810. .8929,9409. .9514)
                     /locus_tag="LL22NC03-121E8.1-001"
     mRNA            join(7219. .8044,8241. .8279,8810. .8929,9409. .9514)
                     /locus_tag="LL22NC03-121E8.1-001"
     CDS             join(7598. .8044,8241. .8279,8810. .8866)
                     /locus_tag="LL22NC03-121E8.1-001"
                     /standard_name="OTTHUMP00000028824"
                     /codon_start=1
                     /protein_id="CAI18780.1"
                     /db_xref="GI:56208181"
                     /db_xref="UniProt/TrEMBL:Q5R3S2"
                     /translation="MPPPCTRHGAASAAYSPGNHPHAKHRRGKGSLPLDCVQHGCRDT
                     DKHIPWQRHGPGWLELQQLCVVTELPNEGDGWPRFTVCPEHVHGAWEDQGLTMPGSGF
                     PHSEASLLVGSMAKHGGRGWGSSVSDFSQAARPLGAGQENRNSILEQKMEPSTSREPT
                     AQRWTLRDRWLTRMLKYVLS"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 39372;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15817 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 15868


RESULT 62
HSN121E8/c
LOCUS       HSN121E8               39372 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone LL22NC03-121E8 on chromosome 22,
            complete sequence.
ACCESSION   Z78421
VERSION     Z78421.1  GI:1495467
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 39372)
  AUTHORS   Kettleborough,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep LL22NC03-121E8
            is from the human chromosome 22-specific cosmid library LL22NC03,
            constructed at the Biomedical Sciences Division, Lawrence Livermore
            National Laboratory, Livermore, CA 94550 under the auspices of the
            National Laboratory Gene Library Project sponsored by the US
            Department of Energy.  The source of the flow sorted chromosomes
            was a human/hamster hybrid containing chromosomes Y, 22 and 9.
            VECTOR: lawrist16
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .39372
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /clone="LL22NC03-121E8"
                     /clone_lib="LL22NC03"
     gene            complement(join(5697. .5931,6975. .7281))
                     /locus_tag="LL22NC03-121E8.2-002"
     mRNA            complement(join(5697. .5931,6975. .7281))
                     /locus_tag="LL22NC03-121E8.2-002"
     gene            join(7219. .8044,8241. .8279,8810. .8929,9409. .9514)
                     /locus_tag="LL22NC03-121E8.1-001"
     mRNA            join(7219. .8044,8241. .8279,8810. .8929,9409. .9514)
                     /locus_tag="LL22NC03-121E8.1-001"
     CDS             join(7598. .8044,8241. .8279,8810. .8866)
                     /locus_tag="LL22NC03-121E8.1-001"
                     /standard_name="OTTHUMP00000028824"
                     /codon_start=1
                     /protein_id="CAI18780.1"
                     /db_xref="GI:56208181"
                     /db_xref="UniProt/TrEMBL:Q5R3S2"
                     /translation="MPPPCTRHGAASAAYSPGNHPHAKHRRGKGSLPLDCVQHGCRDT
                     DKHIPWQRHGPGWLELQQLCVVTELPNEGDGWPRFTVCPEHVHGAWEDQGLTMPGSGF
                     PHSEASLLVGSMAKHGGRGWGSSVSDFSQAARPLGAGQENRNSILEQKMEPSTSREPT
                     AQRWTLRDRWLTRMLKYVLS"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 39372;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17381 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 17330


RESULT 63
AL445257/c
LOCUS       AL445257               39596 bp    DNA     linear   HTG 13-JUN-2001
DEFINITION  Homo sapiens chromosome 1 clone RP5-1175N1, 14 unordered pieces.
ACCESSION   AL445257
VERSION     AL445257.1  GI:10716509
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
```

```
            Center project name: dJ1175N1
            ------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 28716 bases at least Q40
            Consensus quality: 32676 bases at least Q30
            Consensus quality: 35450 bases at least Q20
            Insert size: 38296; sum-of-contigs
            Insert size: 138118; 12.9% error; agarose-fp
            Quality coverage: 1.51x in Q20 bases; sum-of-contigs Quality
            coverage: 0.97x in Q20 bases; agarose-fp
            -------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2942: contig of 2942 bp in length
            *     2943     3042: gap of 100 bp
            *     3043     5573: contig of 2531 bp in length
            *     5574     5673: gap of 100 bp
            *     5674     9075: contig of 3402 bp in length
            *     9076     9175: gap of 100 bp
            *     9176    11187: contig of 2012 bp in length
            *    11188    11287: gap of 100 bp
            *    11288    14656: contig of 3369 bp in length
            *    14657    14756: gap of 100 bp
            *    14757    17428: contig of 2672 bp in length
            *    17429    17528: gap of 100 bp
            *    17529    19964: contig of 2436 bp in length
            *    19965    20064: gap of 100 bp
            *    20065    24184: contig of 4120 bp in length
            *    24185    24284: gap of 100 bp
            *    24285    26468: contig of 2184 bp in length
            *    26469    26568: gap of 100 bp
            *    26569    29667: contig of 3099 bp in length
            *    29668    29767: gap of 100 bp
            *    29768    31926: contig of 2159 bp in length
            *    31927    32026: gap of 100 bp
            *    32027    34705: contig of 2679 bp in length
            *    34706    34805: gap of 100 bp
            *    34806    37009: contig of 2204 bp in length
            *    37010    37109: gap of 100 bp
            *    37110    39596: contig of 2487 bp in length.
FEATURES             Location/Qualifiers
     source          1. .39596
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP5-1175N1"
                     /clone_lib="RPCI-5"
     misc_feature    1. .2942
                     /note="assembly_fragment:00018
                     fragment_chain:1"
     misc_feature    3043. .5573
                     /note="assembly_fragment:00371
                     fragment_chain:1"
     misc_feature    5674. .9075
                     /note="assembly_fragment:00430
                     fragment_chain:2"
     misc_feature    9176. .11187
                     /note="assembly_fragment:00095
                     fragment_chain:2"
     misc_feature    11288. .14656
                     /note="assembly_fragment:00113"
     misc_feature    14757. .17428
                     /note="assembly_fragment:00120"
     misc_feature    17529. .19964
                     /note="assembly_fragment:00198"
     misc_feature    20065. .24184
                     /note="assembly_fragment:00245"
     misc_feature    24285. .26468
                     /note="assembly_fragment:00304"
     misc_feature    26569. .29667
                     /note="assembly_fragment:00328"
     misc_feature    29768. .31926
                     /note="assembly_fragment:00481"
     misc_feature    32027. .34705
                     /note="assembly_fragment:00489"
     misc_feature    34806. .37009
                     /note="assembly_fragment:00517"
     misc_feature    37110. .39596
                     /note="assembly_fragment:00532"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 39596;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12668 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12617


RESULT 64
AL359082
LOCUS       AL359082               41372 bp    DNA     linear   PRI 16-APR-2005
DEFINITION  Human DNA sequence from clone RP4-576K7 on chromosome 1 Contains
            part of the FRAP1 (FK506 binding protein 12-rapamycin associated
            protein 1) gene. ESTs, an STS and GSSs, complete sequence.
ACCESSION   AL359082
VERSION     AL359082.16  GI:14329962
KEYWORDS    HTG; binding protein; FRAP1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 41372)
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 8, 2001 this sequence version replaced gi:13445394.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP4-576K7 is from the library RPCI-4 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
```

```
        IMPORTANT: This sequence is not the entire insert of clone
        RP4-576K7 It may be shorter because we sequence overlapping
        sections only once, except for a 100 base overlap.
        The true left end of clone RP4-576K7 is at 1 in this sequence. The
        true left end of clone RP4-635E18 is at 41273 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .41372
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP4-576K7"
                     /clone_lib="RPCI-4"
     gene            join(<401. .540,7465. .7565,10902. .11017,12648. .12725,
                     12938. .13045,18036. .18161,18259. .18390,23230. .23345,
                     24502. .24619,26922. .27170,27861. .27961,28748. .28844,
                     29147. .29245,29573. .29695,29893. .30075,30555. .30689,
                     30903. .31077,33066. .33201,35573. .35720,36331. .36453,
                     40608. .>40690)
                     /gene="FRAP1"
     CDS             join(<401. .540,7465. .7565,10902. .11017,12648. .12725,
                     12938. .13045,18036. .18161,18259. .18390,23230. .23345,
                     24502. .24619,26922. .27170,27861. .27961,28748. .28844,
                     29147. .29245,29573. .29695,29893. .30075,30555. .30689,
                     30903. .31077,33066. .33201,35573. .35720,36331. .36453,
                     40608. .>40690)
                     /gene="FRAP1"
                     /note="match: cDNAs: Em:U11681 Em:U88966 Em:AF152838
                     Em:L34075 Em:L37085 Em:AF046004 Em:AK024393 Em:L35478
                     match: ESTs: Em:BF689449 Em:BG420811 Em:BG174145
                     Em:BE560006 Em:BE560852 Em:AL120924 Em:BF782990
                     Em:BG342777 Em:BE933597 Em:BE871879 Em:AI738514
                     Em:BF039490 Em:AV594302 Em:AU135855 Em:BE911828
                     Em:BE010022 Em:BF118061 Em:R11912 Em:BE163620 Em:BF921354
                     Em:R11865 Em:BG421771 Em:W26223 Em:BF921191 Em:R60442
                     Em:R55558 Em:AA497757 Em:AI614270 Em:BE514616 Em:BE720437
                     Em:T75267 Em:BG250647 Em:BF818028 Em:BF782326 Em:BE236792
                     Em:BG088519 Em:BE835098 Em:F08162 Em:BE892046 Em:T05942
                     Em:AW500990 Em:AV594303 Em:AL079589 Em:BF856122
                     Em:BF920810 Em:BF760012 Em:BF508672
                     match: proteins: Tr:Q9JLN9 Sw:P42346 Tr:Q9Y4I3 Sw:P42345
                     Sw:P32600 Tr:Q9Y7K2 Tr:O14356 Tr:Q9VK45 Tr:O94189
                     Tr:Q9LPM4 Tr:O94188 Sw:P35169 Tr:O01438 Tr:Q9YI30
                     Tr:O76222"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="dJ576K7.1 (FK506 binding protein 12-rapamycin
                     associated protein 1)"
                     /protein_id="CAC42395.1"
                     /db_xref="GI:14530309"
                     /db_xref="GOA:Q96QW8"
                     /db_xref="InterPro:IPR000403"
                     /db_xref="InterPro:IPR003151"
                     /db_xref="InterPro:IPR006025"
                     /db_xref="UniProt/TrEMBL:Q96QW8"
                     /translation="EIQATWYEKLHEWEDALVAYDKKMDTNKDDPELMLGRMRCLEAL
                     GEWGQLHQQCCEKWTLVNDETQAKMARMAAAAAWGLGQWDSMEEYTCMIPRDTHDGAF
                     YRAVLALHQDLFSLAQQCIDKARDLLDAELTAMAGESYSRAYGAMVSCHMLSELEEVI
                     QYKLVPERRRIIRQIWWERLQGCQRIVEDWQKILMVRSLVVSPHEDMRTWLKYASLCG
                     KSGRLALAHKTLVLLLGVDPSRQLDHPLPTVHPQVTYAYMKNMWKSARKIDAFQHMQH
                     FVQTMQQQAQHAIATEDQQHKQELHKLMARCFLKLGEWQLNLQGINESTIPKVLQYYS
                     AATEHDRSWYKAWHAWAVMNFEAVLHYKHQNQARDEKKKLRHASGANITNATTAATTA
                     ATATTTASTEGSNSESEAESTENSPTPSPLQKKVTEDLSKTLLMYTVPAVQGFFRSIS
                     LSRGNNLQDTLRVLTLWFDYGHWPDVNEALVEGVKAIQIDTWLQVIPQLIARIDTPRP
                     LVGRLIHQLLTDIGRYHPQALIYPLTVASKSTTTARHNAANKILKNMCEHSNTLVQQA
                     MMVSEELIRVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKE
                     TSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKQLPQLTSLELQY
                     VSPKLLMCRDLELAVPGTYDPNQPIIRIQSIAPSLQVITSKQRPRKLTLMGSNGHEFV
                     FLLKGHEDLRQDERVMQLFGLVNTLLANDPTSLRKNLSIQRYAVIPLSTNSGLIGWVP
                     HCDTLHALIRDYREKKKILLNIEHRIMLRMAPDYDHLTLMQKVEVFEHAVNNTAGDDL
                     AKLLWLKSPSSEVWFDRRTNYTRSLAVMSMVGYILGLGD"
     repeat_region   942. .1116
                     /note="L2 repeat: matches 2514. .2688 of consensus"
     repeat_region   1403. .1711
                     /note="AluSx repeat: matches 1. .310 of consensus"
     repeat_region   1798. .1919
                     /note="L2 repeat: matches 2588. .2709 of consensus"
     repeat_region   2065. .2274
                     /note="MER39b repeat: matches 193. .386 of consensus"
     repeat_region   2637. .2828
                     /note="AluJo repeat: matches 1. .190 of consensus"
     repeat_region   2829. .2989
                     /note="FAM repeat: matches 1. .174 of consensus"
     repeat_region   2990. .3046
                     /note="AluJo repeat: matches 85. .310 of consensus"
     repeat_region   3302. .3399
                     /note="AluJ/FLAM repeat: matches 2. .99 of consensus"
     repeat_region   3493. .3528
                     /note="4 copies 9 mer tgttgttgt 94% conserved"
     repeat_region   3550. .3592
                     /note="MER3 repeat: matches 169. .209 of consensus"
     repeat_region   3593. .3904
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   3905. .4063
                     /note="MER3 repeat: matches 1. .169 of consensus"
     repeat_region   5278. .5341
                     /note="U6 repeat: matches 1. .65 of consensus"
     repeat_region   6021. .6203
                     /note="MER5A repeat: matches 1. .189 of consensus"
     misc_feature    complement(6653. .7159)
                     /note="match: GSS: Em:AQ686069"
     misc_feature    complement(7139. .7569)
                     /note="match: GSS: Em:AQ801311"
     misc_feature    complement(7462. .7569)
                     /note="match: GSS: Em:AZ380853"
     repeat_region   7877. .8300
                     /note="MER21B repeat: matches 39. .481 of consensus"
     repeat_region   8328. .8426
                     /note="MER21B repeat: matches 568. .669 of consensus"
     misc_feature    8613. .8759
                     /gene="FRAP1"
                     /note="match: GSS: Em:AQ553575"
     repeat_region   8627. .8747
                     /note="AluJb repeat: matches 2. .122 of consensus"
     misc_feature    complement(8749. .9044)
                     /note="match: GSS: Em:AQ531228"
     misc_feature    complement(8750. .9015)
                     /note="match: GSS: Em:AG019010"
     misc_feature    8757. .8963
                     /gene="FRAP1"
                     /note="match: GSS: Em:AQ781128"
     repeat_region   8784. .9073
                     /note="AluSg repeat: matches 1. .292 of consensus"
     misc_feature    complement(8785. .8897)
                     /note="match: GSS: Em:AQ515715"
     misc_feature    8790. .8947
                     /gene="FRAP1"
                     /note="match: GSS: Em:AQ061904"
     repeat_region   9358. .9583
                     /note="MLT1J repeat: matches 235. .463 of consensus"
     repeat_region   9626. .9706
                     /note="L1MC5 repeat: matches 7836. .7916 of consensus"
     repeat_region   11890. .12024
                     /note="AluSq/x repeat: matches 1. .138 of consensus"
     repeat_region   12025. .12313
                     /note="AluSx repeat: matches 1. .297 of consensus"
     repeat_region   13341. .13507
                     /note="MER5B repeat: matches 1. .178 of consensus"
     misc_feature    complement(13761. .14140)
                     /note="match: GSS: Em:AQ089815"
     misc_feature    complement(13852. .14185)
                     /note="match: GSS: Em:AZ030613"
     misc_feature    14226. .14687
                     /gene="FRAP1"
                     /note="match: GSS: Em:AQ760056"
     misc_feature    14226. .14655
```

```
                     /gene="FRAP1"
                     /note="match: GSS: Em:AQ186184"
     repeat_region   15097. .15396
                     /note="AluSx repeat: matches 1. .292 of consensus"
     repeat_region   15397. .15460
                     /note="AluS repeat: matches 1. .69 of consensus"
     repeat_region   15663. .15988
                     /note="AluY repeat: matches 1. .311 of consensus"
     repeat_region   16025. .16296

  Query Match             1.7%;  Score 52;  DB 8;  Length 41372;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      35426 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 35477


RESULT 65
AC004602
LOCUS       AC004602               42572 bp    DNA     linear   PRI 22-APR-1998
DEFINITION  Homo sapiens chromosome 19, cosmid F23487, complete sequence.
ACCESSION   AC004602
VERSION     AC004602.1  GI:3075375
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 42572)
  AUTHORS   Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
            Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,
            Phan,H., Velasco,N., Garnes,J., Danganan,L., Poundstone,P.,
            Christensen,M., Georgescu,A., Avila,J., Liu,S., Attix,C.,
            Andreise,T., Trankheim,M., Amico-Keller,G., Coefield,J., Duarte,S.,
            Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A.,
            Montgomery,M., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O.
            and Carrano,A.V.
  TITLE     Sequence analysis of a 1 Mb region in 19p13.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 42572)
  AUTHORS   Lamerdin,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-1998) Joint Genome Institute, Lawrence Livermore
            National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
COMMENT     Map and sequence oriented from p telomere to centromere. Cosmid
            F23487 should overlap cosmid R30335 to the right. There is
            currently an approx. 8 kb sequence gap to the left between F23487
            and P1-29569. Additional chr 19 map and sequence information is
            available at http://www-bio.llnl.gov/genome/genome.html.
FEATURES             Location/Qualifiers
     source          1. .42572
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19p13.3 between D19S883 and MLLT1"
                     /clone="F23487"
                     /cell_line="UV5HL9-5B"
                     /clone_lib="LL19NC02 F chromosome 19-specific cosmid
                     library"
                     /note="Cosmid library constructed at LLNL from flow-sorted
                     chromosomes from hybrid UV5HL9-5B, which carries
                     chromosome 19 as its only human chromosome."
     repeat_region   complement(1. .63)
                     /rpt_family="Alu"
     repeat_region   complement(68. .330)
                     /rpt_family="AluJb"
     repeat_region   complement(333. .605)
                     /rpt_family="AluJb"
     repeat_region   complement(741. .771)
                     /rpt_family="MIR"
     repeat_region   complement(769. .886)
                     /rpt_family="LINE2"
     repeat_region   1087. .1388
                     /rpt_family="AluSx"
     repeat_region   complement(1389. .1663)
                     /rpt_family="MLT1A2"
     repeat_region   1668. .1832
                     /rpt_family="THE1C"
     repeat_region   1833. .2129
                     /rpt_family="AluSx"
     repeat_region   2139. .2343
                     /rpt_family="THE1C"
     repeat_region   complement(2344. .2404)
                     /rpt_family="MLT1A2"
     repeat_region   2405. .2665
                     /rpt_family="AluSx"
     repeat_region   2677. .2977
                     /rpt_family="AluSg"
     repeat_region   complement(2978. .3002)
                     /rpt_family="AT_rich"
     misc_feature    3029. .3131
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 2, quality: good, score: 72.000"
     misc_feature    3166. .3478
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: excellent, score: 91.000"
     repeat_region   complement(3820. .4131)
                     /rpt_family="AluSx"
     repeat_region   complement(4727. .4762)
                     /rpt_family="MIR"
     repeat_region   4763. .5031
                     /rpt_family="AluSx"
     repeat_region   5032. .5055
                     /rpt_family="(CAA)n"
     repeat_region   complement(5154. .5254)
                     /rpt_family="AluJo"
     repeat_region   5256. .5545
                     /rpt_family="AluSx"
     repeat_region   complement(5546. .5719)
                     /rpt_family="AluJo"
     repeat_region   complement(6013. .6320)
                     /rpt_family="AluSx"
     repeat_region   complement(6339. .6503)
                     /rpt_family="AluSg"
     repeat_region   complement(6504. .6804)
                     /rpt_family="AluSq"
     repeat_region   complement(6806. .6942)
                     /rpt_family="AluSg"
     repeat_region   complement(6945. .7248)
                     /rpt_family="AluJo"
     repeat_region   complement(7754. .8053)
                     /rpt_family="AluSg"
     misc_feature    7955. .8378
                     /note="DDS similarity to H80195 yu56f07.r1 Homo sapiens
                     cDNA clone 230149 5' similar to contains Alu repetitive
                     element;contains HGR repetitive element ;. Score: 788
                     Identity: 419/428 (97%)."
     repeat_region   complement(8082. .8198)
                     /rpt_family="FLAM_C"
     repeat_region   complement(8272. .8312)
                     /rpt_family="MIR"
     gene            8632. .10218
                     /gene="CAPS"
     misc_feature    8632. .8651
                     /gene="CAPS"
                     /note="DDS similarity to R59818 yh11c08.r1 Homo sapiens
                     cDNA clone 42992 5' similar to SP:A49020 A49020
                     CALCIUM-BINDING PROTEIN R2D5; (1. .35); 100% identity."
     misc_feature    8649. .8831
                     /gene="CAPS"
                     /standard_name="exon"
                     /note="DDS similarity to overlspping ESTs:
```

(8649. .8831) AA128281 zl29e12.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 503374 5' similar to TR:E245872 E245872 CALCYPHOSINE; (1. .183); 99% identity.~(8693. .8831) AA359592 EST68631 Fetal lung II Homo sapiens cDNA 5' end similar to calcyphosine; (1. .139); 99% identity.~(8727. .8831) R59818 yh11c08.r1 Homo sapiens cDNA clone 42992 5' similar to SP:A49020 A49020 CALCIUM-BINDING PROTEIN R2D5; (36. .140); 99% identity."

misc_feature 8727. .8831
/gene="CAPS"
/note="predicted exon, program: grail2exons_human_1.3, frame: 0, quality: excellent, score: 100.000"

CDS join(8749. .8831,8905. .9082,9282. .9488,9563. .9664)
/gene="CAPS"
/note="CALCYPHOSINE (CALCIUM BINDING PROTEIN); RD25"
/codon_start=1
/product="CAYP_HUMAN"
/protein_id="AAC14484.1"
/db_xref="GI:3075376"
/translation="MDAVDATMEKLRAQCLSRGASGIQGLARFFRQLDRDGSRSLDAD
EFRQGLAKLGLVLDQAEAEGVCRKWDRNGSGTLDLEEFLRALRPPMSQAREAVIAAAF
AKLDRSGDGVVTVDDLRGVYSGRAHPKVRSGEWTEDEVLRRFLDNFDSSEKDGQVTLA
EFQDYYSGVSASMNTDEEFVAMMTSAWQL"

misc_feature 8905. .9082
/gene="CAPS"
/standard_name="exon"
/note="DDS similarity to overlapping ESTs:
(8905. .9082) AA128281 zl29e12.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 503374 5' similar to TR:E245872 E245872 CALCYPHOSINE; (184. .359); 98% identity.~(8905. .9082) R59818 yh11c08.r1 Homo sapiens cDNA clone 42992 5' similar to SP:A49020 A49020 CALCIUM-BINDING PROTEIN R2D5; (141. .321); 93% identity.~(8905. .9013) AA359592 EST68631 Fetal lung II Homo sapiens cDNA 5' end similar to calcyphosine; (140. .251); 95% identity.~predicted exon, program: grail2exons_human_1.3, frame: 1, quality: good, score: 71.000"

misc_feature 9282. .9488
/gene="CAPS"
/standard_name="exon"
/note="predicted exon, program: grail2exons_human_1.3, frame: 2, quality: excellent, score: 100.000~(9282. .9424) DDS similarity to AA128281 zl29e12.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 503374 5' similar to TR:E245872 E245872 CALCYPHOSINE; (360. .502); 96% identity."

misc_feature 9523. .9832
/gene="CAPS"
/standard_name="exon"
/note="DDS similarity to R88929 yp97h05.r1 Homo sapiens cDNA clone 195417 5' similar to SP:CAYP_CANFA P10463 CALCYPHOSIN ; (19. .359); 99% identity."

misc_feature 9563. .9660
/gene="CAPS"
/note="predicted exon, program: grail2exons_human_1.3, frame: 1, quality: excellent, score: 100.000"

misc_feature complement(9722. .10218)
/note="DDS similarity to overlapiing ESTs:
(10142. .9722) AA470903 ne20h12.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:881831. Score: 807 Identity: 414/422 (98%).

(10147. .9727) AA642553 nq73a11.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1157948. Score: 817 Identity: 416/420 (99%).~(10218. .9777) AA125808 zl29e12.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 503374 3'. Score: 851 Identity: 435/441 (98%)~(10219. .9803) AA483290 ne39b09.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:899705 similar to TR:G191151 G191151 PRO-ALPHA-1 TYPE V COLLAGEN. Score: 804 Identity: 412/415 (99%).~Other EST matches:
AA909760"

misc_feature 10481. .11538
/note="DDS similarity to overlapping ESTs:

Query Match 1.7%; Score 52; DB 8; Length 42572;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 31346 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 31397

RESULT 66
AC016627
LOCUS AC016627 44449 bp DNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens chromosome 19 clone LLNLR-251F1, complete sequence.
ACCESSION AC016627
VERSION AC016627.6 GI:10337634
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 44449)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 44449)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 44449)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Sep 29, 2000 this sequence version replaced gi:7711581.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.3.
FEATURES Location/Qualifiers
source 1. .44449
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="LLNLR-251F1"
ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 44449;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9282 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 9333

RESULT 67
CR936927
LOCUS CR936927 44783 bp DNA linear PRI 03-MAR-2005
DEFINITION Human DNA sequence from clone DAMC-32J6 on chromosome 6, complete sequence.
ACCESSION CR936927
VERSION CR936927.1 GI:60495271
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 44783)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAMC-32J6 is from the DNA-Arts.org BAC library MCF.1 VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .44783
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMC-32J6"
                     /clone_lib="DNA-Arts.org BAC library MCF.1"
ORIGIN

  Query Match              1.7%;  Score 52;  DB 8;  Length 44783;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 190


RESULT 68
AC006139/c
LOCUS       AC006139               46201 bp    DNA     linear   PRI 05-DEC-1998
DEFINITION  Homo sapiens clone UWGC:y55c068 from 6p21, complete sequence.
ACCESSION   AC006139
VERSION     AC006139.1  GI:3970957
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 46201)
  AUTHORS   Janer,M., Guillaudeux,T., Vu,Q., Kutyavin,T., Harter,H. and
            Geraghty,D.E.
  TITLE     Large scale sequence analysis of the human MHC class I region
  JOURNAL   Unpublished (1998)
  REMARK    Fred Hutchinson Cancer Research Center
            The Clinical Research Division
            1100  Fairview Ave. N., P.O. Box 19024
            Seattle, WA 98109-1024
REFERENCE   2  (bases 1 to 46201)
  AUTHORS   Geraghty,D.E. and Olson,M.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-1998) Human Genome Center, University of
            Washington, Box 352145, Seattle, WA 98195, USA
  REMARK    University of Washington Human Genome Center
             Box 352145, Seattle, WA 98195
             Contact: Daniel E. Geraghty (geraghty@fhcrc.org)
COMMENT     Overlapping Sequences:
               5': UWGC:y55c025 (Genbank Accession: AC004209)
               3': UWGC:y55c138 (Genbank Accession: AC004206)
            --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.

             Double stranded (DS) coverage:  84.7%
             DS or two chemistry coverage:   98.6%
             Single stranded regions:           4
            --------------------
            Sequence Validation:
             This sequence has been validated by Multiple Complete Digest
             Mapping. Comparison of the experimentally derived map digest
             fragments with sequence-predicted fragments is given below.
             Small fragments below a variable cutoff (approximately 400-600bp)
             are not mapped and hence do not appear in the table. There are no
             significant remaining discrepancies between the experimental and
             predicted values. Uniquely ordered fragment groups are separated
             by dashed lines.
                   BglII            HindIII            NsiI
                Map      Seq      Map      Seq      Map      Seq
             -------- --------  -------- --------  -------- --------
              3065.60  3075.00   7741.77  7605.00  18011.21 17666.00
             -------- --------  -------- --------  -------- --------
               760.21   760.00   1925.00  1905.00   1487.24  1502.00
             -------- --------  -------- --------  -------- --------
              3367.22  3318.00    551.96   542.00  17901.11 17315.00
             -------- --------  -------- --------  -------- --------
              2061.62  2038.00    550.36   572.00
             -------- --------  -------- --------
              8479.38  8246.00    549.54   526.00
             -------- --------  -------- --------
             12862.98 12605.00  17067.47 16692.00
             -------- --------  -------- --------
               948.71   946.00   1109.30  1099.00
             -------- --------           --------
              5284.45  5123.00    908.39   885.00
             -------- --------  -------- --------
                                 8007.86  7729.00
                                -------- --------
                                 6403.96  6095.00
                                -------- --------.
FEATURES             Location/Qualifiers
     source          1. .46201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21"
                     /clone="CGM1:D123C12"
```

```
                     /sub_clone="UWGC:y55c068"
                     /cell_line="CGM1"
                     /clone_lib="Wash U YAC library"
     repeat_region   143. .751
                     /rpt_family="Alu"
     repeat_region   1012. .1311
                     /rpt_family="Alu"
     repeat_region   1720. .2002
                     /rpt_family="Alu"
     repeat_region   complement(2189. .2306)
                     /rpt_family="LTR8"
     repeat_region   2390. .2642
                     /rpt_family="Alu"
     repeat_region   2823. .3142
                     /rpt_family="Alu"
     repeat_region   complement(3342. .3488)
                     /rpt_family="LTR8"
     repeat_region   complement(3794. .4088)
                     /rpt_family="Alu"
     repeat_region   complement(4597. .4976)
                     /rpt_family="Alu"
     repeat_region   complement(5537. .5816)
                     /rpt_family="Alu"
     repeat_region   complement(6041. .6333)
                     /rpt_family="Alu"
     repeat_region   complement(6479. .12641)
                     /rpt_family="L1"
     repeat_region   complement(11144. .11222)
                     /rpt_family="MER25"
     repeat_region   complement(11442. .11610)
                     /rpt_family="MER25"
     repeat_region   13441. .13727
                     /rpt_family="Alu"
     repeat_region   14018. .14195
                     /rpt_family="Alu"
     repeat_region   complement(14307. .15266)
                     /rpt_family="LTR12"
     repeat_region   15725. .16027
                     /rpt_family="Alu"
     repeat_region   16086. .16350
                     /rpt_family="Alu"
     repeat_region   16386. .16701
                     /rpt_family="Alu"
     repeat_region   17427. .17724
                     /rpt_family="Alu"
     repeat_region   complement(18377. .19150)
                     /rpt_family="MER4"
     repeat_region   complement(18685. .18976)
                     /rpt_family="Alu"
     repeat_region   complement(19202. .19442)
                     /rpt_family="Alu"
     repeat_region   19789. .19923
                     /rpt_family="Alu"
     repeat_region   complement(20024. .20180)
                     /rpt_family="Alu"
     repeat_region   complement(20422. .20679)
                     /rpt_family="Alu"
     repeat_region   complement(20804. .21089)
                     /rpt_family="Alu"
     repeat_region   complement(21281. .21568)
                     /rpt_family="Alu"
     repeat_region   complement(21913. .22032)
                     /rpt_family="Alu"
     repeat_region   complement(22541. .22790)
                     /rpt_family="MER42"
     repeat_region   complement(22816. .22968)
                     /rpt_family="Alu"
     repeat_region   23481. .23767
                     /rpt_family="Alu"
     repeat_region   complement(25929. .26198)
                     /rpt_family="Alu"
     repeat_region   26298. .27083
                     /rpt_family="Alu"
     repeat_region   complement(27761. .28381)
                     /rpt_family="Alu"
     repeat_region   28631. .28906
                     /rpt_family="Alu"
     repeat_region   complement(29405. .29699)
                     /rpt_family="Alu"
     repeat_region   29951. .30236
                     /rpt_family="Alu"
     repeat_region   31046. .31644
                     /rpt_family="Alu"
     repeat_region   33178. .33466
                     /rpt_family="Alu"
     repeat_region   33713. .34001
                     /rpt_family="Alu"
     repeat_region   complement(34116. .34209)
                     /rpt_family="Alu"
     repeat_region   complement(34226. .34369)
                     /rpt_family="Alu"
     repeat_region   34841. .35164
                     /rpt_family="LTR5"
     repeat_region   35170. .35277
                     /rpt_family="LTR5"
     repeat_region   complement(36065. .36319)
                     /rpt_family="Alu"
     repeat_region   37201. .37670
                     /rpt_family="MER1"
     repeat_region   complement(37323. .37418)
                     /rpt_family="MER1"
     repeat_region   38132. .38421
                     /rpt_family="Alu"
     repeat_region   complement(39377. .39729)
                     /rpt_family="MSTAR"
     repeat_region   complement(40127. .40416)
                     /rpt_family="Alu"
     repeat_region   41331. .41643
                     /rpt_family="Alu"
     repeat_region   44180. .44436
                     /rpt_family="Alu"
     repeat_region   complement(44598. .44804)
                     /rpt_family="Alu"
     repeat_region   44890. .45211
                     /rpt_family="Alu"

  Query Match             1.7%;  Score 52;  DB 8;  Length 46201;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20873 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 20822


RESULT 69
AC108747/c
LOCUS       AC108747               46509 bp    DNA     linear   PRI 21-JUN-2002
DEFINITION  Homo sapiens 3 BAC RP11-268E23 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC108747
VERSION     AC108747.5  GI:21535890
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 46509)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
```

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 46509)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (31-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 46509)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (03-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 46509)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (21-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jun 21, 2002 this sequence version replaced gi:21306532.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES Location/Qualifiers

source	1. .46509
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-268E23"
misc_feature	1. .2005
	/note="overlaps bases 182198. .184202 of clone AC008249"
	/function="clone overlap"
repeat_region	complement(4. .320)
	/rpt_family="L1M4"
repeat_region	complement(321. .613)
	/rpt_family="AluSx"
repeat_region	complement(614. .1226)
	/rpt_family="L1M4"
repeat_region	complement(1240. .1444)
	/rpt_family="AluSq"
repeat_region	complement(1457. .1660)
	/rpt_family="L1M4"
repeat_region	complement(1661. .1967)
	/rpt_family="AluY"
repeat_region	complement(1968. .2614)
	/rpt_family="L1M4"
repeat_region	complement(2988. .3470)
	/rpt_family="L1M4"
repeat_region	complement(3471. .3578)
	/rpt_family="AluY"
repeat_region	complement(3589. .4303)
	/rpt_family="L1M4"
repeat_region	complement(4330. .4501)
	/rpt_family="L1MC4"
repeat_region	complement(4564. .5853)
	/rpt_family="L1ME2"
repeat_region	complement(5816. .7283)
	/rpt_family="L1MEc"
repeat_region	complement(7906. .8217)
	/rpt_family="AluSp"
repeat_region	9117. .9352
	/rpt_family="7SK"
repeat_region	9534. .9793
	/rpt_family="L2"
STS	9763. .10087
	/standard_name="177147"
repeat_region	complement(11575. .11871)
	/rpt_family="MER33"
repeat_region	11887. .12036
	/rpt_family="MER5A"
repeat_region	complement(12070. .12165)
	/rpt_family="MER5A"
repeat_region	12208. .12235
	/rpt_family="AT_rich"
repeat_region	16886. .17058
	/rpt_family="(TA)n"

```
     repeat_region   18116. .18138
                     /rpt_family="AT_rich"
     repeat_region   18713. .18768
                     /rpt_family="(TATATG)n"
     repeat_region   18770. .18835
                     /rpt_family="GA-rich"
     repeat_region   19136. .19424
                     /rpt_family="AluSc"
     repeat_region   complement(20236. .20362)
                     /rpt_family="MIR"
     repeat_region   21144. .21183
                     /rpt_family="(CA)n"
     repeat_region   21974. .22010
                     /rpt_family="(CA)n"
     repeat_region   22512. .22559
                     /rpt_family="(TG)n"
     repeat_region   complement(22584. .22635)
                     /rpt_family="MER31B"
     repeat_region   24689. .24727
                     /rpt_family="MER102"
     repeat_region   24728. .25024
                     /rpt_family="AluSx"
     repeat_region   25025. .25240
                     /rpt_family="MER102"
     STS             25577. .25763
                     /standard_name="60878"
     STS             26576. .26677
                     /standard_name="62312"
     repeat_region   26649. .26742
                     /rpt_family="MIR"
     repeat_region   27097. .27117
                     /rpt_family="AT_rich"
     repeat_region   complement(27489. .27606)
                     /rpt_family="MER5A"
     repeat_region   27758. .27935

  Query Match             1.7%;  Score 52;  DB 8;  Length 46509;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45643 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 45592

RESULT 70
CR759912
LOCUS       CR759912               47462 bp    DNA     linear   PRI 16-OCT-2004
DEFINITION  Human DNA sequence from clone DAAP-394O23 on chromosome 6, complete
            sequence.
ACCESSION   CR759912
VERSION     CR759912.6  GI:54262068
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 47462)
  AUTHORS   Gray,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 16, 2004 this sequence version replaced gi:54021829.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAAP-394O23 is from a APD human bac - B Lymphoblastoid Cell Line
            library VECTOR: pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .47462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAAP-394O23"
                     /clone_lib="APD"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 47462;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5661 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 5712


RESULT 71
AL390837
LOCUS       AL390837               51090 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP5-1168A16 map p36.2, 5 unordered
            pieces.
ACCESSION   AL390837
VERSION     AL390837.7  GI:13785281
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plumb,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 24, 2001 this sequence version replaced gi:13750888.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ1168A16
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
```

```
            Consensus quality: 47312 bases at least Q40
            Consensus quality: 47821 bases at least Q30
            Consensus quality: 48428 bases at least Q20
            Insert size: 50690; sum-of-contigs
            Insert size: 85481; 31.1% error; agarose-fp
            Quality coverage: 6.45x in Q20 bases; sum-of-contigs Quality
            coverage: 4.61x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    34808: contig of 34808 bp in length
            *    34809    34908: gap of 100 bp
            *    34909    37152: contig of 2244 bp in length
            *    37153    37252: gap of 100 bp
            *    37253    39725: contig of 2473 bp in length
            *    39726    39825: gap of 100 bp
            *    39826    45246: contig of 5421 bp in length
            *    45247    45346: gap of 100 bp
            *    45347    51090: contig of 5744 bp in length.
FEATURES             Location/Qualifiers
     source          1. .51090
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p36.2"
                     /clone="RP5-1168A16"
                     /clone_lib="RPCI-5"
     misc_feature    1. .34808
                     /note="assembly_fragment:00349"
     misc_feature    34909. .37152
                     /note="assembly_fragment:00417"
     misc_feature    37253. .39725
                     /note="assembly_fragment:01180"
     misc_feature    39826. .45246
                     /note="assembly_fragment:00548
                     fragment_chain:1
                     clone_end:T7
                     vector_side:right"
     misc_feature    45347. .51090
                     /note="assembly_fragment:00789
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 51090;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32300 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 32351


RESULT 72
AC103989
LOCUS       AC103989               59744 bp    DNA     linear   HTG 01-DEC-2001
DEFINITION  Homo sapiens chromosome 18 clone RP11-407C18 map 18, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC103989
VERSION     AC103989.1  GI:17223277
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 59744)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-407C18
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 59744)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L21885
                Center clone name: 407_C_18
            --------------.
            * NOTE: This record contains 75 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      677: contig of 677 bp in length
            *      678      777: gap of 100 bp
            *      778     1448: contig of 671 bp in length
            *     1449     1548: gap of 100 bp
            *     1549     2244: contig of 696 bp in length
            *     2245     2344: gap of 100 bp
            *     2345     3032: contig of 688 bp in length
            *     3033     3132: gap of 100 bp
            *     3133     3820: contig of 688 bp in length
            *     3821     3920: gap of 100 bp
            *     3921     4605: contig of 685 bp in length
            *     4606     4705: gap of 100 bp
            *     4706     5411: contig of 706 bp in length
            *     5412     5511: gap of 100 bp
            *     5512     6210: contig of 699 bp in length
            *     6211     6310: gap of 100 bp
            *     6311     7016: contig of 706 bp in length
            *     7017     7116: gap of 100 bp
            *     7117     7837: contig of 721 bp in length
```

```
*     7838     7937: gap of 100 bp
*     7938     8620: contig of 683 bp in length
*     8621     8720: gap of 100 bp
*     8721     9421: contig of 701 bp in length
*     9422     9521: gap of 100 bp
*     9522    10227: contig of 706 bp in length
*    10228    10327: gap of 100 bp
*    10328    11026: contig of 699 bp in length
*    11027    11126: gap of 100 bp
*    11127    11821: contig of 695 bp in length
*    11822    11921: gap of 100 bp
*    11922    12614: contig of 693 bp in length
*    12615    12714: gap of 100 bp
*    12715    13404: contig of 690 bp in length
*    13405    13504: gap of 100 bp
*    13505    14209: contig of 705 bp in length
*    14210    14309: gap of 100 bp
*    14310    14988: contig of 679 bp in length
*    14989    15088: gap of 100 bp
*    15089    15793: contig of 705 bp in length
*    15794    15893: gap of 100 bp
*    15894    16524: contig of 631 bp in length
*    16525    16624: gap of 100 bp
*    16625    17326: contig of 702 bp in length
*    17327    17426: gap of 100 bp
*    17427    18118: contig of 692 bp in length
*    18119    18218: gap of 100 bp
*    18219    18917: contig of 699 bp in length
*    18918    19017: gap of 100 bp
*    19018    19711: contig of 694 bp in length
*    19712    19811: gap of 100 bp
*    19812    20510: contig of 699 bp in length
*    20511    20610: gap of 100 bp
*    20611    21289: contig of 679 bp in length
*    21290    21389: gap of 100 bp
*    21390    22091: contig of 702 bp in length
*    22092    22191: gap of 100 bp
*    22192    22899: contig of 708 bp in length
*    22900    22999: gap of 100 bp
*    23000    23695: contig of 696 bp in length
*    23696    23795: gap of 100 bp
*    23796    24494: contig of 699 bp in length
*    24495    24594: gap of 100 bp
*    24595    25287: contig of 693 bp in length
*    25288    25387: gap of 100 bp
*    25388    26065: contig of 678 bp in length
*    26066    26165: gap of 100 bp
*    26166    26868: contig of 703 bp in length
*    26869    26968: gap of 100 bp
*    26969    27734: contig of 766 bp in length
*    27735    27834: gap of 100 bp
*    27835    28629: contig of 795 bp in length
*    28630    28729: gap of 100 bp
*    28730    29432: contig of 703 bp in length
*    29433    29532: gap of 100 bp
*    29533    30211: contig of 679 bp in length
*    30212    30311: gap of 100 bp
*    30312    31011: contig of 700 bp in length
*    31012    31111: gap of 100 bp
*    31112    31800: contig of 689 bp in length
*    31801    31900: gap of 100 bp
*    31901    32595: contig of 695 bp in length
*    32596    32695: gap of 100 bp
*    32696    33386: contig of 691 bp in length
*    33387    33486: gap of 100 bp
*    33487    34177: contig of 691 bp in length
*    34178    34277: gap of 100 bp
*    34278    34944: contig of 667 bp in length
*    34945    35044: gap of 100 bp
*    35045    35745: contig of 701 bp in length
*    35746    35845: gap of 100 bp
*    35846    36549: contig of 704 bp in length
*    36550    36649: gap of 100 bp
*    36650    37364: contig of 715 bp in length
*    37365    37464: gap of 100 bp
*    37465    38174: contig of 710 bp in length
*    38175    38274: gap of 100 bp
*    38275    38976: contig of 702 bp in length
*    38977    39076: gap of 100 bp
*    39077    39778: contig of 702 bp in length
*    39779    39878: gap of 100 bp
*    39879    40564: contig of 686 bp in length
*    40565    40664: gap of 100 bp
*    40665    41354: contig of 690 bp in length
*    41355    41454: gap of 100 bp
*    41455    42144: contig of 690 bp in length
*    42145    42244: gap of 100 bp
*    42245    42951: contig of 707 bp in length
*    42952    43051: gap of 100 bp
*    43052    43759: contig of 708 bp in length
*    43760    43859: gap of 100 bp
*    43860    44560: contig of 701 bp in length
*    44561    44660: gap of 100 bp
*    44661    45366: contig of 706 bp in length
*    45367    45466: gap of 100 bp
*    45467    46142: contig of 676 bp in length
*    46143    46242: gap of 100 bp
*    46243    46952: contig of 710 bp in length
*    46953    47052: gap of 100 bp
*    47053    47750: contig of 698 bp in length
*    47751    47850: gap of 100 bp
*    47851    48546: contig of 696 bp in length
*    48547    48646: gap of 100 bp
*    48647    49342: contig of 696 bp in length
*    49343    49442: gap of 100 bp
*    49443    50120: contig of 678 bp in length
*    50121    50220: gap of 100 bp
*    50221    50926: contig of 706 bp in length
*    50927    51026: gap of 100 bp
*    51027    51714: contig of 688 bp in length
*    51715    51814: gap of 100 bp
*    51815    52518: contig of 704 bp in length
*    52519    52618: gap of 100 bp
*    52619    53329: contig of 711 bp in length
*    53330    53429: gap of 100 bp
*    53430    54114: contig of 685 bp in length
*    54115    54214: gap of 100 bp

  Query Match             1.7%;  Score 52;  DB 14;  Length 59744;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38530 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38581


RESULT 73
AL499606/c
LOCUS       AL499606              62761 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-13J16 on chromosome 6 Contains
            the gene for winged helix/forkhead transcription factor (HFH1) and
            two CpG islands, complete sequence.
ACCESSION   AL499606
VERSION     AL499606.18  GI:17973944
KEYWORDS    HTG; CpG island; HFH1; winged helix/forkhead transcription factor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 62761)
  AUTHORS   Corby,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
```

```
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 20, 2001 this sequence version replaced gi:17902904.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-13J16 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .62761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-13J16"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-13J16"
     misc_feature    4171
                     /note="Clone_right_end: RP1-116B8"
     gene            16713. .18427
                     /gene="FOXQ1"
                     /locus_tag="RP11-13J16.1-001"
     mRNA            16713. .18427
                     /gene="FOXQ1"
                     /locus_tag="RP11-13J16.1-001"
                     /product="forkhead box Q1"
                     /note="match: cDNAs: L13201.1"
     CDS             16945. .18156
                     /gene="FOXQ1"
                     /locus_tag="RP11-13J16.1-001"
                     /standard_name="OTTHUMP00000017774"
                     /note="match: proteins: O54743 O60548 O70220 P56260 Q00939
                     Q12947 Q13461 Q60987 Q61345 Q94H60 Q98937 Q9C009 Q9JJ18
                     Q9JLN7 Q9NS06 Q9PRJ8 Q9QY14"
                     /codon_start=1
                     /product="forkhead box Q1"
                     /protein_id="CAH70503.1"
                     /db_xref="GI:55665392"
                     /db_xref="GOA:Q5VV73"
                     /db_xref="InterPro:IPR001766"
                     /db_xref="UniProt/TrEMBL:Q5VV73"
                     /translation="MKLEVFVPRAAHGDKQGSDLEGAGGSDAPSPLSAAGDDSLGSDG
                     DCAANSPAAGGGARDTQGDGEQSAGGGPGAEEAIPAAAAAAAVVAEGAEAGAAGPGAGG
                     AGSGEGARSKPYTRRPKPPYSYIALIAMAIRDSAGGRLTLAEINEYLMGKFPFFRGSY
                     TGWRNSVRHNLSLNDCFVKVLRDPSRPWGKDNYWMLNPNSEYTFADGVFRRRRKRLSH
                     RAPVPAPGLRPEEAPGLPAAPPPAPAAPASPRMRSPARQEERASPAGKFSSSFAIDSI
                     LRKPFRSRRLRDTAPGTTLQWGAAPCPPLPAFPALLPAAPCRALLPLCAYGAGEPARL
                     GAREAEVPPTAPPLLLAPLPAAAPAKPLRGPAAGGAHLYCPLRLPAALQAASVRRPGP
                     HLPYPVETLLA"
     misc_feature    17171. .17234
                     /gene="FOXQ1"
                     /locus_tag="RP11-13J16.1-001"
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     misc_feature    17634. .17641
                     /gene="FOXQ1"
                     /locus_tag="RP11-13J16.1-001"
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     misc_feature    21165. .21540
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Assembly consistent with restriction digest data."
     misc_feature    21466
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 430bp by restriction digest data."
     misc_feature    23824. .24045
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     misc_feature    60092
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 120bp by restriction digest data."
     misc_feature    60762
                     /note="Clone_left_end: RP4-668J24"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 62761;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37151 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 37100


RESULT 74
AC092080
LOCUS       AC092080               64056 bp    DNA     linear   PRI 18-DEC-2001
DEFINITION  Homo sapiens chromosome 19 clone CTD-2094A2, complete sequence.
ACCESSION   AC092080
VERSION     AC092080.2  GI:17921218
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 64056)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 64056)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 64056)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Dec 18, 2001 this sequence version replaced gi:14495365.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 127.1kb). It is clipped at the overlap with AC010614.
            The number of bases overlapped is 23573.
FEATURES             Location/Qualifiers
```

```
     source          1. .64056
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2094A2"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 64056;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38283 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38334


RESULT 75
AF411057/c
LOCUS       AF411057               66887 bp    DNA     linear   PRI 13-DEC-2001
DEFINITION  Homo sapiens clone BAC_22606 CD28 antigen (CD28) gene, exons 1, 2
            and 3, and partial cds.
ACCESSION   AF411057
VERSION     AF411057.1  GI:17646225
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 66887)
  AUTHORS   Ling,V., Wu,P.W., Finnerty,H.F., Agostino,M.J., Graham,J.R.,
            Chen,S., Jussiff,J.M., Fisk,G.J., Miller,C.P. and Collins,M.
  TITLE     Assembly and Annotation of Human Chromosome 2q33 Sequence
            Containing the CD28, CTLA4, and ICOS Gene Cluster: Analysis by
            Computational, Comparative, and Microarray Approaches
  JOURNAL   Genomics 78 (3), 155-168 (2001)
   PUBMED   11735222
REFERENCE   2  (bases 1 to 66887)
  AUTHORS   Ling,V., Wu,P.W., Finnerty,H.F., Agostino,M.J., Graham,J.R.,
            Chen,S., Jussiff,J., Fisk,G.J., Miller,C.P. and Collins,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) Department of Immunology, Genetics
            Institute/Wyeth Research, 200 Cambridge Park Drive, Cambridge, MA
            02140, USA
FEATURES             Location/Qualifiers
     source          1. .66887
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2q33"
                     /clone="BAC_22606"
     misc_feature    7838. .8329
                     /note="similar to NADH:ubiquinone oxidoreductase"
     gene            42348. .>66887
                     /gene="CD28"
     mRNA            join(42348. .42621,62505. .62861,65540. .>65664)
                     /gene="CD28"
                     /product="CD28 antigen"
     exon            42348. .42621
                     /gene="CD28"
                     /number=1
     5'UTR           42348. .42569
                     /gene="CD28"
     CDS             join(42570. .42621,62505. .62861,65540. .>65664)
                     /gene="CD28"
                     /codon_start=1
                     /product="CD28 antigen"
                     /protein_id="AAL40931.1"
                     /db_xref="GI:17646226"
                     /translation="MLRLLLALNLFPSIQVTGNKILVKQSPMLVAYDNAVNLSCKYSY
                     NLFSREFRASLHKGLDSAVEVCVVYGNYSQQLQVYSKTGFNCDGKLGNESVTFYLQNL
                     YVNQTDIYFCKIEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLFPGPSKPFWVLVVVG
                     GVLACYSLLVTVAFIIFW"
     exon            62505. .62861
                     /gene="CD28"
                     /number=2
     exon            65540. .65664
                     /gene="CD28"
                     /number=3
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 66887;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13733 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 13682


RESULT 76
HSJ735P11/c
LOCUS       HSJ735P11              68959 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP4-735P11 on chromosome Xp11.3-11.4
            Contains STSs and GSSs, complete sequence.
ACCESSION   AL049735
VERSION     AL049735.4  GI:5596978
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 68959)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 27, 1999 this sequence version replaced gi:5101820.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome X, constructed by the Sanger Centre Chromosome X Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            RP4-735P11 is from the library RPCI-4 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .68959
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704P11735"
                     /db_xref="taxon:9606"
                     /map="p11.3-11.4"
                     /clone="RP4-735P11"
```

```
                     /clone_lib="RPCI-4"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 68959;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18500 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 18449


RESULT 77
AC073488/c
LOCUS       AC073488               70393 bp    DNA     linear   PRI 31-DEC-2002
DEFINITION  Homo sapiens X BAC RP11-126O22 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC073488
VERSION     AC073488.23  GI:27436742
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 70393)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 70393)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 70393)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 70393)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 31, 2002 this sequence version replaced gi:21629140.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
            ml.
FEATURES             Location/Qualifiers
     source          1. .70393
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="RP11-126O22"
     repeat_region   132. .259
                     /rpt_family="7SLRNA"
     repeat_region   260. .556
                     /rpt_family="AluSq"
     misc_feature    463. .10771
                     /note="sized by restriction digest and formal approved"
                     /function="unresolved tandem repeat"
     repeat_region   558. .597
                     /rpt_family="(TAGA)n"
     repeat_region   598. .765
                     /rpt_family="7SLRNA"
     repeat_region   766. .789
                     /rpt_family="AT_rich"
     repeat_region   2233. .2309
                     /rpt_family="5S"
     repeat_region   2480. .2584
```

```
                /rpt_family="5S"
     repeat_region   4455. .4504
                /rpt_family="5S"
     repeat_region   9224. .9273
                /rpt_family="5S"
     repeat_region   11202. .11248
                /rpt_family="5S"
     repeat_region   12246. .12295
                /rpt_family="5S"
     repeat_region   13041. .13077
                /rpt_family="(TAAAA)n"
     repeat_region   13850. .13876
                /rpt_family="AT_rich"
     repeat_region   14808. .14893
                /rpt_family="GA-rich"
     repeat_region   15573. .15596
                /rpt_family="AT_rich"
     repeat_region   complement(15634. .15894)
                /rpt_family="HAL1"
     repeat_region   complement(15895. .16191)
                /rpt_family="AluJb"
     repeat_region   complement(16192. .16332)
                /rpt_family="HAL1"
     repeat_region   complement(16511. .16685)
                /rpt_family="HAL1"
     repeat_region   complement(16911. .17203)
                /rpt_family="AluJb"
     repeat_region   complement(17215. .17526)
                /rpt_family="AluY"
     repeat_region   complement(17535. .17678)
                /rpt_family="Charlie5"
     repeat_region   18289. .18430
                /rpt_family="FLAM_C"
     repeat_region   18591. .18662
                /rpt_family="CT-rich"
     repeat_region   complement(18777. .18925)
                /rpt_family="MIR"
     repeat_region   20368. .20668
                /rpt_family="AluY"
     repeat_region   21055. .21187
                /rpt_family="FLAM_C"
     repeat_region   21192. .21508
                /rpt_family="AluJo"
     repeat_region   complement(21517. .21531)
                /rpt_family="AluJo"
     repeat_region   21532. .21553
                /rpt_family="(TTTA)n"
     repeat_region   complement(21554. .21704)
                /rpt_family="AluJo"
     repeat_region   21705. .21727
                /rpt_family="(TTTA)n"
     repeat_region   complement(21728. .21847)
                /rpt_family="AluJo"
     repeat_region   complement(21866. .22592)
                /rpt_family="L1MB5"
     repeat_region   23263. .23400
                /rpt_family="AluJo"
     repeat_region   23405. .23688
                /rpt_family="AluJb"
     repeat_region   23740. .24344
                /rpt_family="L1ME1"
     repeat_region   24346. .24589
                /rpt_family="MLT1D"
     repeat_region   complement(24590. .25032)
                /rpt_family="MSTB"
     repeat_region   25033. .25256
                /rpt_family="MLT1D"
     repeat_region   25258. .25377

Query Match             1.7%;  Score 52;  DB 8;  Length 70393;
Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52328 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 52277


RESULT 78
CR388372
LOCUS       CR388372               70448 bp    DNA     linear   PRI 19-JUL-2004
DEFINITION  Human DNA sequence from clone DADB-87H23 on chromosome 6, complete
            sequence.
ACCESSION   CR388372
VERSION     CR388372.6  GI:50399819
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 70448)
  AUTHORS   Wood,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 19, 2004 this sequence version replaced gi:49614041.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            DADB-87H23 is from a DNA-arts DBB human bac library VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .70448
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DADB-87H23"
                     /clone_lib="DNA-arts-BAC.1-DBB.1"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 70448;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7753 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7804
```

```
RESULT 79
AC109598
LOCUS       AC109598               73597 bp    DNA     linear   HTG 06-FEB-2002
DEFINITION  Homo sapiens chromosome 16 clone CTD-2589A16, WORKING DRAFT
            SEQUENCE, 27 unordered pieces.
ACCESSION   AC109598
VERSION     AC109598.1  GI:18542986
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 73597)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 73597)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 824173
            Center clone name: CITB-E1_2589A16
            -------
            Summary Statistics
            Consensus quality: 49396 bases at least Q40
            Consensus quality: 55879 bases at least Q30
            Consensus quality: 59662 bases at least Q20
            Estimated insert size: 0; null estimation
            Estimated insert size: 70997; sum-of-contigs estimation
            Quality coverage:  2.1474836E7 in Q20 bases; null estimation
            Quality coverage:  20.98 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 27 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1010: contig of 1010 bp in length
            *     1011     1110: gap of unknown length
            *     1111     2146: contig of 1036 bp in length
            *     2147     2246: gap of unknown length
            *     2247     3280: contig of 1034 bp in length
            *     3281     3380: gap of unknown length
            *     3381     4759: contig of 1379 bp in length
            *     4760     4859: gap of unknown length
            *     4860     6014: contig of 1155 bp in length
            *     6015     6114: gap of unknown length
            *     6115     7375: contig of 1261 bp in length
            *     7376     7475: gap of unknown length
            *     7476     8931: contig of 1456 bp in length
            *     8932     9031: gap of unknown length
            *     9032    10428: contig of 1397 bp in length
            *    10429    10528: gap of unknown length
            *    10529    11943: contig of 1415 bp in length
            *    11944    12043: gap of unknown length
            *    12044    13282: contig of 1239 bp in length
            *    13283    13382: gap of unknown length
            *    13383    14589: contig of 1207 bp in length
            *    14590    14689: gap of unknown length
            *    14690    15744: contig of 1055 bp in length
            *    15745    15844: gap of unknown length
            *    15845    17600: contig of 1756 bp in length
            *    17601    17700: gap of unknown length
            *    17701    19026: contig of 1326 bp in length
            *    19027    19126: gap of unknown length
            *    19127    20180: contig of 1054 bp in length
            *    20181    20280: gap of unknown length
            *    20281    21726: contig of 1446 bp in length
            *    21727    21826: gap of unknown length
            *    21827    23387: contig of 1561 bp in length
            *    23388    23487: gap of unknown length
            *    23488    25327: contig of 1840 bp in length
            *    25328    25427: gap of unknown length
            *    25428    26682: contig of 1255 bp in length
            *    26683    26782: gap of unknown length
            *    26783    29073: contig of 2291 bp in length
            *    29074    29173: gap of unknown length
            *    29174    30458: contig of 1285 bp in length
            *    30459    30558: gap of unknown length
            *    30559    32899: contig of 2341 bp in length
            *    32900    32999: gap of unknown length
            *    33000    34752: contig of 1753 bp in length
            *    34753    34852: gap of unknown length
            *    34853    36765: contig of 1913 bp in length
            *    36766    36865: gap of unknown length
            *    36866    38523: contig of 1658 bp in length
            *    38524    38623: gap of unknown length
            *    38624    40041: contig of 1418 bp in length
            *    40042    40141: gap of unknown length
            *    40142    73597: contig of 33456 bp in length.
FEATURES             Location/Qualifiers
     source          1. .73597
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-2589A16"
                     /clone_lib="CalTech human BAC library D"
     gap             1011. .1110
                     /estimated_length=unknown
     gap             2147. .2246
                     /estimated_length=unknown
     gap             3281. .3380
                     /estimated_length=unknown
     gap             4760. .4859
                     /estimated_length=unknown
     gap             6015. .6114
                     /estimated_length=unknown
     gap             7376. .7475
                     /estimated_length=unknown
     gap             8932. .9031
                     /estimated_length=unknown
     gap             10429. .10528
                     /estimated_length=unknown
     gap             11944. .12043
                     /estimated_length=unknown
     gap             13283. .13382
                     /estimated_length=unknown
     gap             14590. .14689
                     /estimated_length=unknown
     gap             15745. .15844
                     /estimated_length=unknown
     gap             17601. .17700
                     /estimated_length=unknown
     gap             19027. .19126
                     /estimated_length=unknown
     gap             20181. .20280
                     /estimated_length=unknown
     gap             21727. .21826
                     /estimated_length=unknown
     gap             23388. .23487
                     /estimated_length=unknown
     gap             25328. .25427
                     /estimated_length=unknown
```

```
     gap             26683. .26782
                     /estimated_length=unknown
     gap             29074. .29173
                     /estimated_length=unknown
     gap             30459. .30558
                     /estimated_length=unknown
     gap             32900. .32999
                     /estimated_length=unknown
     gap             34753. .34852
                     /estimated_length=unknown
     gap             36766. .36865
                     /estimated_length=unknown
     gap             38524. .38623
                     /estimated_length=unknown
     gap             40042. .40141
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 73597;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71157 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 71208


RESULT 80
AL391561/c
LOCUS       AL391561               75525 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-99P18 on chromosome 1 Contains
            part of the FRAP1 gene for FK506 binding protein 12-rapamycin
            associated protein 1, a novel gene and the 5' end of the gene for
            angiopoietin-like factor (CDT6), complete sequence.
ACCESSION   AL391561
VERSION     AL391561.20  GI:15808186
KEYWORDS    HTG; CDT6; FK506; FRAP1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 75525)
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 28, 2001 this sequence version replaced gi:15626141.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-99P18 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .75525
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-99P18"
                     /clone_lib="RPCI-11.1"
     misc_feature    2000
                     /note="Clone_right_end: RP4-635E18"
     gene            join(complement(AL049653.7:74499. .74561),
                     complement(AL049653.7:71302. .71477),
                     complement(AL049653.7:70539. .70647),
                     complement(AL049653.7:68987. .69219),
                     complement(AL049653.7:68046. .68246),
                     complement(AL049653.7:65893. .66027),
                     complement(AL049653.7:59873. .60148),
                     complement(AL049653.7:59679. .59787),
                     complement(AL049653.7:55168. .55354),
                     complement(AL049653.7:53607. .53735),
                     complement(AL049653.7:52357. .52601),
                     complement(AL049653.7:50456. .50671),
                     complement(AL049653.7:49897. .50102),
                     complement(AL049653.7:46197. .46319),
                     complement(AL049653.7:45452. .45541),
                     complement(AL049653.7:44490. .44582),
                     complement(AL049653.7:43354. .43488),
                     complement(AL049653.7:42979. .43108),
                     complement(AL049653.7:40722. .40972),
                     complement(AL049653.7:28202. .28288),
                     complement(AL049653.7:25453. .25620),
                     complement(AL049653.7:24850. .24962),
                     complement(AL049653.7:24366. .24528),
                     complement(AL049653.7:22868. .22960),
                     complement(AL049653.7:21366. .21512),
                     complement(AL049653.7:16615. .16757),
                     complement(AL049653.7:11595. .11757),
                     complement(AL049653.7:11312. .11457),
                     complement(53021. .53096),complement(42731. .42870),
                     complement(35705. .35805),complement(32255. .32370),
                     complement(30547. .30624),complement(30227. .30334),
                     complement(25112. .25237),complement(24883. .25014),
                     complement(19930. .20045),complement(18659. .18776),
                     complement(16108. .16356),complement(15317. .15417),
                     complement(14434. .14530),complement(14033. .14131),
                     complement(13583. .13705),complement(13203. .13385),
                     complement(12589. .12723),complement(12201. .12375),
                     complement(10077. .10212),complement(7558. .7705),
                     complement(6825. .6947),complement(2583. .2665),
                     complement(AL109811.40:111744. .111816),
                     complement(AL109811.40:111161. .111235),
                     complement(AL109811.40:110666. .110801),
                     complement(AL109811.40:109200. .109265),
                     complement(AL109811.40:105997. .106077),
                     complement(AL109811.40:105638. .105718),
                     complement(AL109811.40:104529. .104634),
                     complement(AL109811.40:102883. .103848))
                     /gene="FRAP1"
                     /locus_tag="RP4-635E18.1-001"
     mRNA            join(complement(AL049653.7:74499. .74561),
                     complement(AL049653.7:71302. .71477),
                     complement(AL049653.7:70539. .70647),
                     complement(AL049653.7:68987. .69219),
                     complement(AL049653.7:68046. .68246),
                     complement(AL049653.7:65893. .66027),
                     complement(AL049653.7:59873. .60148),
                     complement(AL049653.7:59679. .59787),
                     complement(AL049653.7:55168. .55354),
                     complement(AL049653.7:53607. .53735),
                     complement(AL049653.7:52357. .52601),
                     complement(AL049653.7:50456. .50671),
```

```
                     complement(AL049653.7:49897. .50102),
                     complement(AL049653.7:46197. .46319),
                     complement(AL049653.7:45452. .45541),
                     complement(AL049653.7:44490. .44582),
                     complement(AL049653.7:43354. .43488),
                     complement(AL049653.7:42979. .43108),
                     complement(AL049653.7:40722. .40972),
                     complement(AL049653.7:28202. .28288),
                     complement(AL049653.7:25453. .25620),
                     complement(AL049653.7:24850. .24962),
                     complement(AL049653.7:24366. .24528),
                     complement(AL049653.7:22868. .22960),
                     complement(AL049653.7:21366. .21512),
                     complement(AL049653.7:16615. .16757),
                     complement(AL049653.7:11595. .11757),
                     complement(AL049653.7:11312. .11457),
                     complement(53021. .53096),complement(42731. .42870),
                     complement(35705. .35805),complement(32255. .32370),
                     complement(30547. .30624),complement(30227. .30334),
                     complement(25112. .25237),complement(24883. .25014),
                     complement(19930. .20045),complement(18659. .18776),
                     complement(16108. .16356),complement(15317. .15417),
                     complement(14434. .14530),complement(14033. .14131),
                     complement(13583. .13705),complement(13203. .13385),
                     complement(12589. .12723),complement(12201. .12375),
                     complement(10077. .10212),complement(7558. .7705),
                     complement(6825. .6947),complement(2583. .2665),
                     complement(AL109811.40:111744. .111816),
                     complement(AL109811.40:111161. .111235),
                     complement(AL109811.40:110666. .110801),
                     complement(AL109811.40:109200. .109265),
                     complement(AL109811.40:105997. .106077),
                     complement(AL109811.40:105638. .105718),
                     complement(AL109811.40:104529. .104634),
                     complement(AL109811.40:102883. .103848))
                     /gene="FRAP1"
                     /locus_tag="RP4-635E18.1-001"
                     /product="FK506 binding protein 12-rapamycin associated
                     protein 1"
                     /note="match: ESTs: Em:AA262510.1 Em:AA282830.1
                     Em:AA725390.1 Em:AA737024.1 Em:AA767875.1 Em:AA768596.1
                     Em:AL044764.1 Em:AL047770.1 Em:AL079658.1 Em:AL120924.1
                     Em:AL704254.1 Em:AU098639.1 Em:AU135855.1 Em:AW500990.1
                     Em:AW960244.1 Em:AW976130.1 Em:BE163620.1 Em:BE539676.1
                     Em:BE560852.1 Em:BE871879.1 Em:BE884193.1 Em:BE898570.1
                     Em:BF689449.1 Em:BF761404.1 Em:BF761418.1 Em:BF871677.1
                     Em:BF893441.1 Em:BG250647.1 Em:BG420811.1 Em:BG421771.1
                     Em:BG768920.1 Em:BG900350.1 Em:BI090093.1 Em:BI258757.1
                     Em:BI460098.1 Em:BM148769.1 Em:BM149445.1 Em:BM150508.1
                     Em:BM151339.1 Em:BM153242.1 Em:BM193414.1 Em:BM461878.1
                     Em:BM470895.1 Em:BM473815.1 Em:BM721256.1 Em:BM724353.1
                     Em:BM799652.1 Em:BM999881.1 Em:BQ010420.1 Em:BQ024837.1
                     Em:BQ065337.1 Em:BQ073602.1 Em:BQ212803.1 Em:BQ219542.1
                     Em:BQ221375.1 Em:BQ440695.1 Em:BQ440903.1 Em:BU165455.1
                     Em:BU431325.1 Em:BU507234.1 Em:BU630556.1 Em:BU741782.1
                     Em:BU932965.1 Em:BX090957.1 Em:BX437773.1 Em:BX476615.1
                     Em:BX476627.1 Em:BX481180.1 Em:BX508363.1 Em:CA406561.1
                     Em:CA429532.1 Em:CA432283.1 Em:CB130997.1 Em:CB426435.1
                     Em:CD367026.1 Em:CD370823.1 Em:CD540746.1 Em:CD557696.1
                     Em:CD654263.1 Em:CD657588.1 Em:CD700609.1 Em:T05942.1
                     match: cDNAs: Em:AF152838.1 Em:L34075.1 Em:U88966.1"
CDS                  join(complement(AL049653.7:71302. .71463),
                     complement(AL049653.7:70539. .70647),
                     complement(AL049653.7:68987. .69219),
                     complement(AL049653.7:68046. .68246),
                     complement(AL049653.7:65893. .66027),
                     complement(AL049653.7:59873. .60148),
                     complement(AL049653.7:59679. .59787),
                     complement(AL049653.7:55168. .55354),
                     complement(AL049653.7:53607. .53735),
                     complement(AL049653.7:52357. .52601),
                     complement(AL049653.7:50456. .50671),
                     complement(AL049653.7:49897. .50102),
                     complement(AL049653.7:46197. .46319),
                     complement(AL049653.7:45452. .45541),
                     complement(AL049653.7:44490. .44582),
                     complement(AL049653.7:43354. .43488),
                     complement(AL049653.7:42979. .43108),
                     complement(AL049653.7:40722. .40972),
                     complement(AL049653.7:28202. .28288),
                     complement(AL049653.7:25453. .25620),
                     complement(AL049653.7:24850. .24962),
                     complement(AL049653.7:24366. .24528),

  Query Match             1.7%;  Score 52;  DB 8;  Length 75525;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7852 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7801


RESULT 81
AC055813
LOCUS       AC055813               80117 bp    DNA     linear   PRI 22-FEB-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-209M4, complete sequence.
ACCESSION   AC055813
VERSION     AC055813.9  GI:18860756
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 80117)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-209M4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 80117)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 80117)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
```

```
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 80117)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 22, 2002 this sequence version replaced gi:15055348.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7949
                Center clone name: 209_M_4
            --------------

            Only the last 80.1 kilobases of this clone are being submitted.
            The remainder overlaps accession number AC004149 [WICGR project
            L302].
FEATURES             Location/Qualifiers
     source          1. .80117
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-209M4"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(146. .434)
                     /rpt_family="AluSg"
     repeat_region   complement(438. .655)
                     /rpt_family="MIR"
     repeat_region   complement(974. .1129)
                     /rpt_family="MER20"
     repeat_region   complement(1270. .1438)
                     /rpt_family="AluJo"
     repeat_region   1439. .1465
                     /rpt_family="(TTTTA)n"
     repeat_region   complement(1466. .1584)
                     /rpt_family="AluJo"
     repeat_region   complement(1785. .1933)
                     /rpt_family="MER5B"
     repeat_region   2339. .2487
                     /rpt_family="MIR"
     repeat_region   complement(2492. .2787)
                     /rpt_family="AluSg"
     repeat_region   complement(2796. .2902)
                     /rpt_family="MIR"
     repeat_region   complement(2903. .3165)
                     /rpt_family="AluSq"
     repeat_region   complement(3166. .3193)
                     /rpt_family="MIR"
     repeat_region   3572. .3644
                     /rpt_family="MIR"
     repeat_region   4524. .4583
                     /rpt_family="MIR3"
     repeat_region   complement(4755. .4962)
                     /rpt_family="MIR"
     repeat_region   6397. .6738
                     /rpt_family="L2"
     repeat_region   complement(6766. .6875)
                     /rpt_family="MLT1H"
     repeat_region   7202. .7270
                     /rpt_family="L2"
     repeat_region   7739. .7825
                     /rpt_family="L1MC4"
     repeat_region   complement(7826. .8122)
                     /rpt_family="AluSc"
     repeat_region   8123. .8268
                     /rpt_family="L1MC4"
     repeat_region   8269. .8568
                     /rpt_family="AluJo"
     repeat_region   8569. .8631
                     /rpt_family="L1MC4"
     repeat_region   complement(9151. .9305)
                     /rpt_family="MIR"
     repeat_region   9365. .9390
                     /rpt_family="AT_rich"
     repeat_region   complement(9405. .9708)
                     /rpt_family="AluY"
     repeat_region   complement(9709. .9969)
                     /rpt_family="AluJo"
     repeat_region   10197. .10458
                     /rpt_family="MLT1K"
     repeat_region   10758. .10895
                     /rpt_family="MIR"
     repeat_region   12676. .12773
                     /rpt_family="MIR"
     repeat_region   complement(12784. .12997)
                     /rpt_family="MIR"
     repeat_region   complement(13220. .13513)
                     /rpt_family="AluJo"
     repeat_region   13627. .13787
                     /rpt_family="L2"
     repeat_region   complement(13821. .13953)
                     /rpt_family="MIR"
     repeat_region   14746. .14899
                     /rpt_family="MIR"
     repeat_region   complement(14947. .15144)
                     /rpt_family="MIR"
     repeat_region   15520. .15668
                     /rpt_family="AluJo"
     repeat_region   15669. .15960
                     /rpt_family="AluSx"
     repeat_region   15961. .16128
                     /rpt_family="AluJo"
     repeat_region   16208. .16400
                     /rpt_family="MIR"
```

```
     repeat_region   16454. .16586
                     /rpt_family="L2"
     repeat_region   16707. .16783

  Query Match             1.7%;  Score 52;  DB 8;  Length 80117;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54323 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 54374


RESULT 82
BX927220
LOCUS       BX927220               82724 bp    DNA     linear   PRI 09-APR-2004
DEFINITION  Human DNA sequence from clone DAMA-199F12 on chromosome 6, complete
            sequence.
ACCESSION   BX927220
VERSION     BX927220.8  GI:46357732
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 82724)
  AUTHORS   Wood,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 9, 2004 this sequence version replaced gi:46016496.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAMA-199F12 is from the DNA-Arts human BAC library MANN.1 VECTOR:
            pBeloBAC11.
FEATURES                Location/Qualifiers
     source             1. .82724
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /clone="DAMA-199F12"
                        /clone_lib="DNA-Arts BAC library MANN.1"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 82724;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        705 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 756


RESULT 83
AR242855
LOCUS       AR242855               92139 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 1 from patent US 6475739.
ACCESSION   AR242855
VERSION     AR242855.1  GI:27289517
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 92139)
  AUTHORS   Brunkow,M.E., Proll,S., Paeper,B. and Staehling-Hampton,K.
  TITLE     Methods for identifying genomic deletions
  JOURNAL   Patent: US 6475739-A 1 05-NOV-2002;
            Celltech R&D, Inc.; Bothell, WA
FEATURES                Location/Qualifiers
     source             1. .92139
                        /organism="unknown"
                        /mol_type="genomic DNA"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 92139;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85294 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85345


RESULT 84
AX384907
LOCUS       AX384907               92139 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0210455.
ACCESSION   AX384907
VERSION     AX384907.1  GI:19578035
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Brunkow,M.E., Proll,S. and Paeper,B.
  TITLE     Methods for identifying genomic deletions
  JOURNAL   Patent: WO 0210455-A 1 07-FEB-2002;
            Celltech R & D, Inc. (US) ; Straehling-Hampton, Karen (US)
FEATURES                Location/Qualifiers
     source             1. .92139
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 92139;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85294 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85345
```

```
RESULT 85
CR751233/c
LOCUS       CR751233               92442 bp    DNA     linear   HTG 20-AUG-2004
DEFINITION  Homo sapiens chromosome 6 clone DASS-94P17, WORKING DRAFT SEQUENCE,
            2 unordered pieces.
ACCESSION   CR751233
VERSION     CR751233.1  GI:51491315
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 92442)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ------------- Project Information
            Center project name: bSS94P17
            ------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 91963 bases at least Q40
            Consensus quality: 92179 bases at least Q30
            Consensus quality: 92265 bases at least Q20
            Insert size: 92342; sum-of-contigs
            Insert size: 99297; 0.8% error; agarose-fp
            Quality coverage: 6.50x in Q20 bases; sum-of-contigs Quality
            coverage: 6.14x in Q20 bases; agarose-fp
            -------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    30199: contig of 30199 bp in length
            *    30200    30299: gap of 100 bp
            *    30300    92442: contig of 62143 bp in length.
FEATURES             Location/Qualifiers
     source          1. .92442
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DASS-94P17"
                     /clone_lib="DNA-arts-BAC.1-SSTO.1"
     misc_feature    1. .30199
                     /note="assembly_fragment:00002.0"
     misc_feature    30300. .92442
                     /note="assembly_fragment:00375.0"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 92442;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54359 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 54308


RESULT 86
AF397423/c
LOCUS       AF397423               93790 bp    DNA     linear   PRI 04-SEP-2002
DEFINITION  Homo sapiens SOST gene, 3' UTR, MEOX1 gene, 5' UTR, and SOST/MEOX1
            intergenic region.
ACCESSION   AF397423
VERSION     AF397423.1  GI:19880618
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 93790)
  AUTHORS   Staehling-Hampton,K., Proll,S., Paeper,B., Zhao,L., Charmley,P.,
            Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
            Papapoulos,S., Hamersma,H. and Brunkow,M.E.
  TITLE     A 52-kb deletion in the SOST-MEOX1 intergenic region on 17q12-q21
            is associated with van Buchem disease in the Dutch population
  JOURNAL   Am. J. Med. Genet. 110 (2), 144-152 (2002)
   PUBMED   12116252
REFERENCE   2  (bases 1 to 93790)
  AUTHORS   Staehling-Hampton,K., Proll,S., Paeper,B., Zhao,L., Charmley,P.,
            Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
            Papapoulos,S., Hamersma,H. and Brunkow,M.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2001) Genomics, Celltech R&D Inc., 1631 220th St
            SE, Bothell, WA 98021, USA
FEATURES             Location/Qualifiers
     source          1. .93790
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q21-q22"
     gene            <1. .1612
                     /gene="SOST"
     mRNA            <1. .1612
                     /gene="SOST"
     3'UTR           1. .1612
                     /gene="SOST"
     STS             12298. .12676
                     /standard_name="D17S951"
     variation       36239. .87958
                     /note="Van Buchem causative deletion"
     STS             79495. .79823
                     /standard_name="D17S1789"
     gene            93762. .>93790
                     /gene="MEOX1"
     mRNA            93762. .>93790
                     /gene="MEOX1"
     5'UTR           93762. .93790
                     /gene="MEOX1"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 93790;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8458 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8407


RESULT 87
AL645729
LOCUS       AL645729               96217 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-318E23 on chromosome 1 Contains
            the GMEB1 gene for glucocorticoid modulatory element binding
            protein 1, the 5' end of the gene for high-glucose-regulated
            protein 8 (HGRG8) and a CpG island, complete sequence.
ACCESSION   AL645729
VERSION     AL645729.14  GI:21322389
KEYWORDS    HTG; GMEB1; HGRG8.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 96217)
  AUTHORS   Hammond,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 4, 2002 this sequence version replaced gi:20451424.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-318E23 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .96217
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-318E23"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-318E23"
     gene            join(14331. .14438,29182. .29339,35703. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .64952)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-001"
     mRNA            join(14331. .14438,29182. .29339,35703. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .64952)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-001"
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /note="match: ESTs: Em:AI221554.1 Em:AI693203.1
                     Em:CA418589.1
                     match: cDNAs: Em:AK000892.1 Em:AK023126.1"
     gene            join(14340. .14438,35703. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .>59833)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-004"
     mRNA            join(14340. .14438,35703. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .>59833)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-004"
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /note="match: ESTs: Em:BX332399.1"
     gene            join(14377. .14438,29185. .29339,35703. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .60485)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-002"
     mRNA            join(14377. .14438,29185. .29339,35703. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .60485)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-002"
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /note="match: cDNAs: Em:AF173868.1 Em:AF203694.1
                     Em:BC001473.1 Em:BC001473.2"
     gene            join(14379. .14438,29182. .29339,35673. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .60472)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-003"
     mRNA            join(14379. .14438,29182. .29339,35673. .35785,37184. .37308,
                     38541. .38644,42526. .42683,48037. .48168,49791. .49928,
                     56119. .56241,59672. .60472)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-003"
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /note="match: ESTs: Em:AA148980.1 Em:BI464962.1
                     Em:BU177938.1
                     match: cDNAs: Em:AF099013.1"
     CDS             join(14383. .14438,35703. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .>59833)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-004"
                     /standard_name="OTTHUMP00000035905"
                     /codon_start=1
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /protein_id="CAH72426.1"
                     /db_xref="GI:55666388"
                     /db_xref="GOA:Q5VSZ8"
                     /db_xref="InterPro:IPR000770"
                     /db_xref="UniProt/TrEMBL:Q5VSZ8"
                     /translation="MLRRLPALAARRPPARRRRIYEAGSENNTAVVAVETHTIHKIEE
                     GIDTGTIEANEDMEIAYPITCGESKAILLWKKFVCPGINVKCVKFNDQLISPKHFVHL
                     AGKSTLKDWKRAIRLGGIMLRKMMDSGQIDFYQHDKVCSNTCRSTKFDLLISSARAPV
                     PGQQTSVVQTPTSADGSITQIAISEESMEEAGLEWNSALTAAVTMATEEGVKKDSEEI
                     SEDTLMFWKGIADVGLMEEVVCNIQKEIEELLRGVQQRLIQAPFQVTDAAVLNNVAHT
                     FGLMDTVKKVLDNRRNQVEQGEEQFLYTLTDLERQLEEQKKQGQDHRLKSQTVQNVVL
                     MPVSTPKPPKRPRLQRPASTTVLSPS"
     misc_feature    14627. .14639
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-004"
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     gene            join(29163. .29339,35673. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .60664)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-005"
     mRNA            join(29163. .29339,35673. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .60664)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-005"
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /note="match: cDNAs: Em:AF210433.1"
     CDS             join(29212. .29339,35673. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .60372)
```

```
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-003"
                     /standard_name="OTTHUMP00000003781"
                     /codon_start=1
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /protein_id="CAH72427.1"
                     /db_xref="GI:55666389"
                     /db_xref="InterPro:IPR000770"
                     /db_xref="InterPro:IPR010919"
                     /translation="MANAEVSVPVGDVVVVPTEGNEGENPEDTKTQVILQLQPVQQGL
                     FIDGHFYNRIYEAGSENNTAVVAVETHTIHKIEEGIDTGTIEANEDMEIAYPITCGES
                     KAILLWKKFVCPGINVKCVKFNDQLISPKHFVHLAGKSTLKDWKRAIRLGGIMLRKMM
                     DSGQIDFYQHDKVCSNTCRSTKFDLLISSARAPVPGQQTSVVQTPTSADGSITQIAIS
                     EESMEEAGLEWNSALTAAVTMATEEGVKKDSEEISEDTLMFWKGIADVGLMEEVVCNI
                     QKEIEELLRGVQQRLIQAPFQVTDAAVLNNVAHTFGLMDTVKKVLDNRRNQVEQGEEQ
                     FLYTLTDLERQLEEQKKQGQDHRLKSQTVQNVVLMPVSTPKPPKRPRLQRPASTTVLS
                     PSPPVQQPQFTVISPITITPVGQSFSMGNIPVATLSQGSSPVTVHTLPSGPQLFRYAT
                     VVSSAKSSSPDTVTIHPSSSLALLSSTAMQDGSTLGNMTTMVSPVELVAMESGLTSAI
                     QAVESTSEDGQTIIEIDPAPDPEAEDTEGKAVILETELRTEEKVVAEMEEHQHQVHNV
                     EIVVLED"
     CDS             join(29212. .29339,35703. .35785,37184. .37308,38541. .38644,
                     42526. .42683,48037. .48168,49791. .49928,56119. .56241,
                     59672. .60372)
                     /gene="GMEB1"
                     /locus_tag="RP11-318E23.1-001"
                     /standard_name="OTTHUMP00000003779"
                     /note="match: proteins: Sw:Q9JL60 Sw:Q9Y692 Tr:Q80Y88"
                     /codon_start=1
                     /product="glucocorticoid modulatory element binding
                     protein 1"
                     /protein_id="CAH72428.1"
                     /db_xref="GI:55666390"
                     /translation="MANAEVSVPVGDVVVVPTEGNEGENPEDTKTQVILQLQPVQQGI
                     YEAGSENNTAVVAVETHTIHKIEEGIDTGTIEANEDMEIAYPITCGESKAILLWKKFV
                     CPGINVKCVKFNDQLISPKHFVHLAGKSTLKDWKRAIRLGGIMLRKMMDSGQIDFYQH
                     DKVCSNTCRSTKFDLLISSARAPVPGQQTSVVQTPTSADGSITQIAISEESMEEAGLE
                     WNSALTAAVTMATEEGVKKDSEEISEDTLMFWKGIADVGLMEEVVCNIQKEIEELLRG
                     VQQRLIQAPFQVTDAAVLNNVAHTFGLMDTVKKVLDNRRNQVEQGEEQFLYTLTDLER
                     QLEEQKKQGQDHRLKSQTVQNVVLMPVSTPKPPKRPRLQRPASTTVLSPSPPVQQPQF
                     TVISPITITPVGQSFSMGNIPVATLSQGSSPVTVHTLPSGPQLFRYATVVSSAKSSSP
                     DTVTIHPSSSLALLSSTAMQDGSTLGNMTTMVSPVELVAMESGLTSAIQAVESTSEDG
                     QTIIEIDPAPDPEAEDTEGKAVILETELRTEEKVVAEMEEHQHQVHNVEIVVLED"
     polyA_site      60511
                     /gene="GMEB1"

  Query Match             1.7%;  Score 52;  DB 8;  Length 96217;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74409 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 74460


RESULT 88
AL844213
LOCUS       AL844213               97630 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone DAQB-352F19 on chromosome 6 Contains
            the TMPOP1 pseudogene for thymopoietin-like 1, the SUCLA2P
            pseudogene for succinate-CoA ligase, ADP-forming, beta subunit, the
            RANP1 pseudogene for RAN, member RAS oncogene family pseudogene 1,
            HLA-E gene for major histocompatibility complex, class I, E, and 6
            CpG islands, complete sequence.
ACCESSION   AL844213
VERSION     AL844213.2  GI:21953190
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 97630)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 24, 2002 this sequence version replaced gi:21912387.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            DAQB-352F19 is from a DNA-arts QBL human bac library VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .97630
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAQB-352F19"
                     /clone_lib="DNA-arts-BAC.1-QBL.1"
     misc_feature    1
                     /note="Clone_left_end: DAQB-352F19"
     gene            join(27698. .28117,28351. .29083)
                     /locus_tag="DAQB-352F19.1-001"
                     /pseudo
     CDS             join(27698. .28117,28351. .29083)
                     /locus_tag="DAQB-352F19.1-001"
                     /note="match: proteins: AAC25390 P42167 Q61030 Q61031"
                     /pseudo
                     /codon_start=1
                     /product="thymopoietin-like 1"
     gene            30105. .31413
                     /locus_tag="DAQB-352F19.2-001"
                     /pseudo
     CDS             30105. .31413
                     /locus_tag="DAQB-352F19.2-001"
                     /note="match: proteins: O82662 O97580 Q9VHJ8 Q9YI37"
                     /pseudo
                     /codon_start=1
                     /product="succinate-CoA ligase, ADP-forming, beta subunit
                     pseudogene"
     gene            47103. .47753
                     /locus_tag="DAQB-352F19.3-001"
                     /pseudo
     CDS             47103. .47753
                     /locus_tag="DAQB-352F19.3-001"
                     /note="match: proteins: AAM15923 P51148 Q9FE79 Q9VNG6"
                     /pseudo
                     /codon_start=1
                     /product="RAN, member RAS oncogene family pseudogene 1"
     misc_feature    49811
                     /note="Clone_right_end: DAQB-39G10"
     gene            50669. .55108
                     /gene="HLA-E"
                     /locus_tag="DAQB-352F19.4-001"
```

```
     mRNA            join(50669. .50755,50886. .51155,51400. .51675,52297. .52572,
                     52697. .52813,53564. .53596,53701. .53743,53909. .55108)
                     /gene="HLA-E"
                     /locus_tag="DAQB-352F19.4-001"
                     /product="major histocompatibility complex, class I, E"
                     /note="match: ESTs: AL545738.1 BM917068.1 BM917589.1
                     BQ052237.1 BQ689518.1
                     match: cDNAs: AK098434.1 BC002578.1 BC004297.1 M20022.1"
     CDS             join(50692. .50755,50886. .51155,51400. .51675,52297. .52572,
                     52697. .52813,53564. .53596,53701. .53741)
                     /gene="HLA-E"
                     /locus_tag="DAQB-352F19.4-001"
                     /standard_name="OTTHUMP00000014549"
                     /note="match: proteins: O46840 Q31521 Q9BT83 Q9MXG3
                     Q9MXQ5"
                     /codon_start=1
                     /product="major histocompatibility complex, class I, E"
                     /protein_id="CAI18596.1"
                     /db_xref="GI:56206333"
                     /db_xref="InterPro:IPR001039"
                     /db_xref="InterPro:IPR003006"
                     /db_xref="InterPro:IPR003597"
                     /db_xref="InterPro:IPR007110"
                     /translation="MVDGTLLLLLSEALALTQTWAGSHSLKYFHTSVSRPGRGEPRFI
                     SVGYVDDTQFVRFDNDAASPRMVPRAPWMEQEGSEYWDRETRSARDTAQIFRVNLRTL
                     RGYYNQSEAGSHTLQWMHGCELGPDRRFLRGYEQFAYDGKDYLTLNEDLRSWTAVDTA
                     AQISEQKSNDASEAEHQRAYLEDTCVEWLHKYLEKGKETLLHLEPPKTHVTHHPISDH
                     EATLRCWALGFYPAEITLTWQQDGEGHTQDTELVETRPAGDGTFQKWAAVVVPSGEEQ
                     RYTCHVQHEGLPEPVTLRWKPASQPTIPIVGIIAGLVLLGSVVSGAVVAAVIWRKKSS
                     GGKGGSYSKAEWSDSAQGSESHSL"
     polyA_signal    55331. .55336
     polyA_site      55354
     misc_feature    97630
                     /note="Clone_right_end: DAQB-352F19"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 97630;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73737 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 73788


RESULT 89
AC093203/c
LOCUS       AC093203              101769 bp    DNA     linear   HTG 07-MAR-2002
DEFINITION  Homo sapiens chromosome 5 clone CTC-262A8, WORKING DRAFT SEQUENCE,
            2 ordered pieces.
ACCESSION   AC093203
VERSION     AC093203.4  GI:19224825
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 101769)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101769)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 101769)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 7, 2002 this sequence version replaced gi:15426059.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 285795
            Center clone name: CIT-HSPC_262A8
            -------
            Summary Statistics
            Consensus quality: 101567 bases at least Q40
            Consensus quality: 101611 bases at least Q30
            Consensus quality: 101641 bases at least Q20
            Estimated insert size: 104000; agarose-fp estimation
            Estimated insert size: 101669; sum-of-contigs estimation
            Quality coverage:  34.54 in Q20 bases; agarose-fp estimation
            Quality coverage:  35.33 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    12059: contig of 12059 bp in length
            *    12060    12159: gap of unknown length
            *    12160   101769: contig of 89610 bp in length.
FEATURES             Location/Qualifiers
     source          1. .101769
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-262A8"
                     /clone_lib="CalTech human BAC library C"
     gap             12060. .12159
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 101769;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      35123 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 35072


RESULT 90
AC005377/c
LOCUS       AC005377              102311 bp    DNA     linear   PRI 03-OCT-2003
DEFINITION  Homo sapiens PAC clone RP5-1136G2 from 7, complete sequence.
ACCESSION   AC005377
VERSION     AC005377.2  GI:4699956
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102311)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 102311)
  AUTHORS   Bauer,C., Ozanich,A. and Tucci,S.
  TITLE     The sequence of Homo sapiens PAC clone RP5-1136G2
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 102311)
  AUTHORS   Waterston,R.H.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 102311)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 102311)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 102311)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 102311)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Apr 28, 1999 this sequence version replaced gi:3907488.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_DJ1136G02
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-5, prepared by
            Pieter de Jong and coworkers at http://www.chori.org using the
            method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
            The library is from one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR:  pCYPAC2

            NEIGHBORING SEQUENCE INFORMATION:
            The  clone sequenced to the right is RP4-659J6, 200 bp overlap.
            Actual start of this clone is at base position 1 of RP5-1136G2
            actual end is at 69895 of DJ0659J06.
FEATURES             Location/Qualifiers
     source          1. .102311
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP5-1136G2"
                     /clone_lib="RPCI-5"
     repeat_region   1. .81
                     /rpt_family="Alu"
     repeat_region   82. .200
                     /rpt_family="L1"
     repeat_region   201. .518
                     /rpt_family="Alu"
     repeat_region   519. .936
                     /rpt_family="L1"
     repeat_region   1437. .1545
                     /rpt_family="MaLR"
     repeat_region   1546. .1590
                     /rpt_family="MIR"
     repeat_region   1591. .1886
                     /rpt_family="Alu"
     repeat_region   1897. .2186
                     /rpt_family="Alu"
     repeat_region   2187. .2259
                     /rpt_family="MIR"
     repeat_region   2281. .2339
                     /rpt_family="CR1"
     repeat_region   2359. .2407
                     /rpt_family="ERVL"
     repeat_region   2425. .2484
                     /rpt_family="(CA)n"
     repeat_region   2840. .2938
                     /rpt_family="L1"
     repeat_region   3563. .3678
                     /rpt_family="L2"
     repeat_region   3679. .3887
                     /rpt_family="Alu"
     repeat_region   3888. .4200
                     /rpt_family="Alu"
     repeat_region   4201. .4328
                     /rpt_family="Alu"
     repeat_region   4334. .4514
                     /rpt_family="L2"
     repeat_region   4633. .4806
                     /rpt_family="L2"
     repeat_region   4808. .5251
                     /rpt_family="MaLR"
     repeat_region   5343. .5499
                     /rpt_family="L2"
     repeat_region   6032. .6161
                     /rpt_family="MIR"
     repeat_region   6740. .6863
                     /rpt_family="MIR"
     repeat_region   7570. .7829
                     /rpt_family="L1"
     repeat_region   8297. .8422
                     /rpt_family="L2"
     repeat_region   8474. .8512
                     /rpt_family="L1"
     repeat_region   8513. .8816
                     /rpt_family="Alu"
     repeat_region   8817. .8894
                     /rpt_family="L1"
     repeat_region   8905. .9046
                     /rpt_family="L1"
     repeat_region   9443. .9533
                     /rpt_family="L1"
     repeat_region   9545. .9594
                     /rpt_family="MaLR"
     repeat_region   9595. .9649
                     /rpt_family="MaLR"
     repeat_region   9650. .9945
                     /rpt_family="Alu"
```

```
     repeat_region   9946. .10330
                     /rpt_family="MaLR"
     repeat_region   10331. .10638
                     /rpt_family="MaLR"
     repeat_region   10639. .10949
                     /rpt_family="Alu"
     repeat_region   10950. .11092
                     /rpt_family="MaLR"
     repeat_region   11100. .11224
                     /rpt_family="L1"
     repeat_region   11381. .11577
                     /rpt_family="L1"
     repeat_region   11578. .11870
                     /rpt_family="Alu"
     repeat_region   11871. .11950
                     /rpt_family="L1"
     repeat_region   12482. .12783
                     /rpt_family="Alu"
     repeat_region   13738. .13897
                     /rpt_family="MER1_type"
     repeat_region   13975. .14102
                     /rpt_family="MIR"
     repeat_region   14199. .14263
                     /rpt_family="L2"
     repeat_region   14313. .14623
                     /rpt_family="Alu"
     repeat_region   14698. .14798
                     /rpt_family="L1"
     repeat_region   14877. .15185
                     /rpt_family="Alu"
     repeat_region   15190. .15283
                     /rpt_family="L1"
     repeat_region   15858. .16008

  Query Match             1.7%;  Score 52;  DB 8;  Length 102311;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75623 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 75572


RESULT 91
AL359699/c
LOCUS       AL359699              105902 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-130F5 on chromosome 1 Contains a
            SLAM family member 6 (SLAMF6) pseudogene, the gene for kinase
            interacting with leukemia-associated gene (KIS) and two CpG
            islands, complete sequence.
ACCESSION   AL359699 AC068536
VERSION     AL359699.12  GI:18476598
KEYWORDS    HTG; KIS; SLAMF6.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 105902)
  AUTHORS   Chapman,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 1, 2002 this sequence version replaced gi:18250758.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-130F5 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            Draft Sequence Produced by Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            http://genome.wustl.edu/gsc/index.shtml
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .105902
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-130F5"
                     /clone_lib="RPCI-11.1"
     misc_feature    2000
                     /note="Clone_right_end: RP11-565P22"
     gene            17885. .18222
                     /locus_tag="RP11-130F5.1-001"
                     /pseudo
     CDS             17885. .18222
                     /locus_tag="RP11-130F5.1-001"
                     /note="match: proteins: Tr:AAQ88526 Tr:Q96DU3 Tr:Q96DV0
                     Tr:Q9ET40"
                     /pseudo
                     /codon_start=1
                     /product="SLAM family member 6 (SLAMF6) pseudogene"
     gene            join(70179. .70604,72291. .72583,73249. .73440,76090. .76184,
                     84836. .84912,85061. .85159,90438. .90526,94740. .101946)
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
     mRNA            join(70179. .70604,72291. .72583,73249. .73440,76090. .76184,
                     84836. .84912,85061. .85159,90438. .90526,94740. .101946)
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
                     /product="U2AF homology motif (UHM) kinase 1"
                     /note="match: cDNAs: Em:AJ536197.1 Em:AK013347.1
                     Em:BC026046.1 Em:BC058732.1 Em:U70372.2 Em:X98374.1
                     Em:Y10725.2"
     CDS             join(70337. .70604,72291. .72583,73249. .73440,76090. .76184,
                     84836. .84912,85061. .85159,90438. .90526,94740. .94886)
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
                     /standard_name="OTTHUMP00000029529"
                     /note="match: proteins: Sw:AAH58732 Sw:P97343 Tr:AAO13515
                     Tr:CAD60192 Tr:Q63285 Tr:Q8TAS1"
                     /codon_start=1
                     /product="U2AF homology motif (UHM) kinase 1"
                     /protein_id="CAH70392.1"
                     /db_xref="GI:55663991"
                     /db_xref="Genew:19683"
                     /db_xref="GOA:Q8TAS1"
                     /db_xref="InterPro:IPR000504"
                     /db_xref="InterPro:IPR000719"
                     /db_xref="InterPro:IPR001245"
                     /db_xref="InterPro:IPR002290"
                     /db_xref="InterPro:IPR008271"
                     /db_xref="InterPro:IPR011009"
```

```
                     /db_xref="UniProt/Swiss-Prot:Q8TAS1"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
     gene            join(70421. .70604,72291. .73440,76090. .76184,84836. .84912,
                     85061. .85159,90438. .90526,94740. .96390)
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-002"
     mRNA            join(70421. .70604,72291. .73440,76090. .76184,84836. .84912,
                     85061. .85159,90438. .90526,94740. .96390)
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-002"
                     /product="U2AF homology motif (UHM) kinase 1"
                     /note="match: cDNAs: Em:AL834136.1"
     misc_feature    96919
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
                     /note="Clone_left_end: RP11-359K18"
     polyA_signal    101941. .101946
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
     polyA_site      101946
                     /gene="UHMK1"
                     /locus_tag="RP11-130F5.2-001"
     misc_feature    105902
                     /note="Clone_right_end: RP11-130F5"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 105902;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15718 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 15667


RESULT 92
AC074136
LOCUS       AC074136               105956 bp    DNA     linear   PRI 03-APR-2003
DEFINITION  Homo sapiens chromosome 16 clone CTD-2318B16, complete sequence.
ACCESSION   AC074136
VERSION     AC074136.4  GI:29501849
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 105956)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 105956)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 105956)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 105956)
  AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 3, 2003 this sequence version replaced gi:22417329.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center and Los Alamos
            National Laboratory
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .105956
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="CTD-2318B16"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 105956;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36426 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 36477


RESULT 93
AF225899/c
LOCUS       AF225899               106539 bp    DNA     linear   PRI 23-FEB-2000
DEFINITION  Homo sapiens PAC clone 219d7, complete sequence.
ACCESSION   AF225899
VERSION     AF225899.1  GI:7021528
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 106539)
  AUTHORS   Fitzpatrick,E.S., Hammond,H.A., DeAngelis,D.M., Soderman,A.R.,
            Wright,J.L., Liu,X., Larson,D., McGowan,J., Ziegler,S.,
            Pritchard,L., Hess,J.F., Todd,J., Caskey,C.T. and Metzker,M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2000) Department of Human Genetics, Merck & Co.,
            Inc., SumneyTown Pike, West Point, PA 19486, USA
FEATURES             Location/Qualifiers
     source          1. .106539
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="PAC 219d7"
     repeat_region   complement(34. .104)
                     /rpt_family="MER5"
     repeat_region   complement(384. .662)
                     /rpt_family="Alu"
     repeat_region   1021. .1304
                     /rpt_family="Alu"
     repeat_region   1800. .2135
                     /rpt_family="Alu"
     repeat_region   2139. .2439
                     /rpt_family="Alu"
     repeat_region   3214. .3294
                     /rpt_family="MLT2B2"
     repeat_region   3588. .3657
                     /rpt_family="MLT2B2"
     repeat_region   3982. .4274
                     /rpt_family="Alu"
     repeat_region   5406. .5550
                     /rpt_family="MER3"
     repeat_region   5444. .5509
                     /rpt_family="MER33"
```

```
repeat_region   complement(5558. .5834)
                /rpt_family="Alu"
repeat_region   5852. .5898
                /rpt_family="MER3"
repeat_region   complement(6534. .6827)
                /rpt_family="Alu"
repeat_region   complement(8239. .8568)
                /rpt_family="LTR12"
repeat_region   11864. .12159
                /rpt_family="Alu"
repeat_region   complement(12838. .13137)
                /rpt_family="Alu"
repeat_region   complement(14450. .14691)
                /rpt_family="Alu"
repeat_region   complement(15004. .15291)
                /rpt_family="Alu"
repeat_region   complement(16143. .16422)
                /rpt_family="Alu"
repeat_region   complement(16813. .17096)
                /rpt_family="Alu"
repeat_region   complement(17528. .17706)
                /rpt_family="MIR"
repeat_region   21610. .21890
                /rpt_family="Alu"
repeat_region   23049. .23329
                /rpt_family="Alu"
repeat_region   24379. .24585
                /rpt_family="L1"
repeat_region   complement(27993. .28279)
                /rpt_family="Alu"
repeat_region   28842. .28933
                /rpt_family="MIR"
STS             30439. .30783
                /note="L18270 Human chromosome 2 STS UT426"
repeat_region   complement(31317. .31610)
                /rpt_family="Alu"
repeat_region   32555. .37857
                /rpt_family="L1"
repeat_region   33985. .34281
                /rpt_family="Alu"
repeat_region   35580. .35863
                /rpt_family="Alu"
repeat_region   complement(37874. .38708)
                /rpt_family="L1"
repeat_region   38616. .38712
                /rpt_family="MIR"
repeat_region   complement(41075. .41331)
                /rpt_family="Alu"
repeat_region   complement(42064. .42184)
                /rpt_family="Alu"
repeat_region   42309. .42413
                /rpt_family="MIR"
repeat_region   complement(42689. .42777)
                /rpt_family="MIR"
repeat_region   43887. .44156
                /rpt_family="Alu"
repeat_region   46963. .47240
                /rpt_family="Alu"
repeat_region   complement(47837. .48028)
                /rpt_family="MIR"
repeat_region   complement(51874. .52142)
                /rpt_family="Alu"
repeat_region   52754. .52801
                /rpt_family="MER2"
repeat_region   58298. .58529
                /rpt_family="Alu"
repeat_region   58600. .58721
                /rpt_family="L1"
repeat_region   complement(59586. .59701)
                /rpt_family="Alu"
repeat_region   60048. .60335
                /rpt_family="Alu"
repeat_region   60964. .61242
                /rpt_family="Alu"
repeat_region   complement(62514. .63136)
                /rpt_family="Alu"
repeat_region   complement(68228. .68344)
                /rpt_family="MER20"
STS             69944. .70696
                /note="G06727 human STS WI-7843"
repeat_region   70776. .71041
                /rpt_family="Alu"
repeat_region   75420. .75585
                /rpt_family="Alu"
repeat_region   complement(77681. .77951)
                /rpt_family="Alu"
repeat_region   complement(78093. .78393)
                /rpt_family="Alu"
repeat_region   79977. .80042
                /rpt_family="THE1"
repeat_region   complement(81069. .81363)
                /rpt_family="Alu"
STS             81561. .81913
                /note="G09915 human STS CHLC.GGAA 19E07.P17317  clone
                GGAA19E07"
repeat_region   complement(82076. .82380)
                /rpt_family="Alu"
repeat_region   82595. .82857
                /rpt_family="Alu"
repeat_region   complement(84782. .85059)
                /rpt_family="Alu"
repeat_region   complement(86131. .86655)
                /rpt_family="Alu"
repeat_region   complement(86698. .86772)
                /rpt_family="MLT1"
repeat_region   87328. .87517
                /rpt_family="Alu"
repeat_region   complement(88025. .88300)
                /rpt_family="Alu"
repeat_region   complement(88382. .88537)
                /rpt_family="Alu"
repeat_region   complement(89103. .92131)
                /rpt_family="L1"
repeat_region   complement(89607. .89870)
                /rpt_family="Alu"
repeat_region   complement(92183. .92468)
                /rpt_family="Alu"
repeat_region   complement(92638. .92723)
                /rpt_family="L1"
repeat_region   complement(93254. .93542)
                /rpt_family="Alu"
repeat_region   93950. .94195
                /rpt_family="MER43"
repeat_region   94819. .94882
                /rpt_family="MER43"
repeat_region   97105. .97472
                /rpt_family="THE1"
STS             99270. .99669
                /note="G08815 human STS CHLC.GGAA19E07.P17317  CLONE
                GGAA19E07"
repeat_region   complement(101199. .101438)
                /rpt_family="Alu"
repeat_region   101753. .102038
                /rpt_family="Alu"
repeat_region   complement(102216. .102638)
                /rpt_family="LOR1"
repeat_region   complement(103792. .104101)
                /rpt_family="L1"
repeat_region   complement(104170. .104465)
                /rpt_family="Alu"
repeat_region   104528. .104642
                /rpt_family="L1"
repeat_region   105260. .105545
                /rpt_family="Alu"
repeat_region   105891. .106161
                /rpt_family="Alu"
```

```
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 106539;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12907 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12856


RESULT 94
BA000025_14/c
WPCOMMENT
  Sequence split into 23 fragments  LOCUS BA000025 Accession BA000025
    Fragment Name      Begin     End
    BA000025_00            1    110000
    BA000025_01       100001    210000
    BA000025_02       200001    310000
    BA000025_03       300001    410000
    BA000025_04       400001    510000
    BA000025_05       500001    610000
    BA000025_06       600001    710000
    BA000025_07       700001    810000
    BA000025_08       800001    910000
    BA000025_09       900001   1010000
    BA000025_10      1000001   1110000
    BA000025_11      1100001   1210000
    BA000025_12      1200001   1310000
    BA000025_13      1300001   1410000
    BA000025_14      1400001   1510000
    BA000025_15      1500001   1610000
    BA000025_16      1600001   1710000
    BA000025_17      1700001   1810000
    BA000025_18      1800001   1910000
    BA000025_19      1900001   2010000
    BA000025_20      2000001   2110000
    BA000025_21      2100001   2210000
    BA000025_22      2200001   2229817
Continuation (15 of 23) of BA000025 from base 1400001 (BA000025 Homo sapiens genomic DNA

  Query Match             1.7%;  Score 52;  DB 8;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29596 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 29545


RESULT 95
BA000041_09/c
WPCOMMENT
  Sequence split into 18 fragments  LOCUS BA000041 Accession BA000041
    Fragment Name      Begin     End
    BA000041_00            1    110000
    BA000041_01       100001    210000
    BA000041_02       200001    310000
    BA000041_03       300001    410000
    BA000041_04       400001    510000
    BA000041_05       500001    610000
    BA000041_06       600001    710000
    BA000041_07       700001    810000
    BA000041_08       800001    910000
    BA000041_09       900001   1010000
    BA000041_10      1000001   1110000
    BA000041_11      1100001   1210000
    BA000041_12      1200001   1310000
    BA000041_13      1300001   1410000
    BA000041_14      1400001   1510000
    BA000041_15      1500001   1610000
    BA000041_16      1600001   1710000
    BA000041_17      1700001   1750601
Continuation (10 of 18) of BA000041 from base 900001 (BA000041 Pan troglodytes DNA, majo

  Query Match             1.7%;  Score 52;  DB 8;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104961 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 104910


RESULT 96
BA000041_10/c
WPCOMMENT
  Sequence split into 18 fragments  LOCUS BA000041 Accession BA000041
    Fragment Name      Begin     End
    BA000041_00            1    110000
    BA000041_01       100001    210000
    BA000041_02       200001    310000
    BA000041_03       300001    410000
    BA000041_04       400001    510000
    BA000041_05       500001    610000
    BA000041_06       600001    710000
    BA000041_07       700001    810000
    BA000041_08       800001    910000
    BA000041_09       900001   1010000
    BA000041_10      1000001   1110000
    BA000041_11      1100001   1210000
    BA000041_12      1200001   1310000
    BA000041_13      1300001   1410000
    BA000041_14      1400001   1510000
    BA000041_15      1500001   1610000
    BA000041_16      1600001   1710000
    BA000041_17      1700001   1750601
Continuation (11 of 18) of BA000041 from base 1000001 (BA000041 Pan troglodytes DNA, maj

  Query Match             1.7%;  Score 52;  DB 8;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4961 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4910


RESULT 97
HS1164I10/c
LOCUS       HS1164I10             110028 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP5-1164I10 on chromosome
            20q13.13-13.2 Contains part of the BIG2 gene for brefeldin
            A-inhibited guanine nucleotide-exchange protein 2, ESTs, STSs, GSSs
            and two CpG islands, complete sequence.
ACCESSION   AL049537
VERSION     AL049537.48  GI:9663111
KEYWORDS    HTG; BIG2; brefeldin A; CpG island; guanine exchanger.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110028)
  AUTHORS   Collier,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 2, 2000 this sequence version replaced gi:9588409.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
```

```
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            RP5-1164I10 is from the library RPCI-5 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .110028
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704I101164"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q13.13-13.2"
                     /clone="RP5-1164I10"
                     /clone_lib="RPCI-5"
     misc_feature    100
                     /note="Clone_right_end: RP5-906C1"
     misc_feature    1871. .1893
                     /note="Tandem repeat. Single clone region. Assembly
                     confirmed by restriction enzyme digest data."
     misc_feature    31598. .31685
                     /note="Single clone region. ALU region. Assembly confirmed
                     by restriction enzyme digest data."
     gene            join(32390. .32510,51699. .51729,52364. .52487,61823. .61969,
                     63205. .63384,64056. .64290,74330. .74398,76372. .76523,
                     79647. .79777,81620. .81854,82826. .82925,83645. .83784,
                     85266. .85374,86516. .86699,95229. .95340,95908. .96113,
                     98804. .98888,98991. .99162,99785. .99936,100056. .100184,
                     101540. .101698,104951. .105098,106283. .106382,
                     108693. .108733,108836. .109005,AL121903.13:5642. .5793,
                     AL121903.13:10804. .10976,AL121903.13:12496. .12656,
                     AL121903.13:14136. .14266,AL121903.13:14403. .14532,
                     AL121903.13:16852. .16987,AL121903.13:17821. .17959,
                     AL121903.13:18068. .18122,AL121903.13:19456. .19570,
                     AL121903.13:23623. .23753,AL121903.13:25885. .26053,
                     AL121903.13:29102. .29240,AL121903.13:32621. .32738,
                     AL121903.13:33595. .37265)
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
     mRNA            join(32390. .32510,51699. .51729,52364. .52487,61823. .61969,
                     63205. .63384,64056. .64290,74330. .74398,76372. .76523,
                     79647. .79777,81620. .81854,82826. .82925,83645. .83784,
                     85266. .85374,86516. .86699,95229. .95340,95908. .96113,
                     98804. .98888,98991. .99162,99785. .99936,100056. .100184,
                     101540. .101698,104951. .105098,106283. .106382,
                     108693. .108733,108836. .109005,AL121903.13:5642. .5793,
                     AL121903.13:10804. .10976,AL121903.13:12496. .12656,
                     AL121903.13:14136. .14266,AL121903.13:14403. .14532,
                     AL121903.13:16852. .16987,AL121903.13:17821. .17959,
                     AL121903.13:18068. .18122,AL121903.13:19456. .19570,
                     AL121903.13:23623. .23753,AL121903.13:25885. .26053,
                     AL121903.13:29102. .29240,AL121903.13:32621. .32738,
                     AL121903.13:33595. .37265)
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /note="match: ESTs: AA157825.1 AA399434.1 AA628490.1
                     AA748620.1 AA765985.1 AI425715.1 AI796261.1 AI813340.1
                     AU124958.1 BB864558.1 BF606564.1 BF739950.1 BF909315.1
                     BI328134.1 BI467887.1 BI687113.1 BI830616.1 BI847308.1
                     BM664817.1 BM826674.1 BM989278.1
                     match: cDNAs: AF084521.1 AK022623.1 AK074303.1"
     CDS             join(32390. .32510,51699. .51729,52364. .52487,61823. .61969,
                     63205. .63384,64056. .64290,74330. .74398,76372. .76523,
                     79647. .79777,81620. .81854,82826. .82925,83645. .83784,
                     85266. .85374,86516. .86699,95229. .95340,95908. .96113,
                     98804. .98888,98991. .99162,99785. .99936,100056. .100184,
                     101540. .101698,104951. .105098,106283. .106382,
                     108693. .108733,108836. .109005,AL121903.13:5642. .5793,
                     AL121903.13:10804. .10976,AL121903.13:12496. .12656,
                     AL121903.13:14136. .14266,AL121903.13:14403. .14532,
                     AL121903.13:16852. .16987,AL121903.13:17821. .17959,
                     AL121903.13:18068. .18122,AL121903.13:19456. .19570,
                     AL121903.13:23623. .23753,AL121903.13:25885. .26053,
                     AL121903.13:29102. .29240,AL121903.13:32621. .32738,
                     AL121903.13:33595. .33771)
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /standard_name="OTTHUMP00000031231"
                     /note="match: proteins: O46382 Q9NKE8 Q9Y6D5 Q9Y6D6"
                     /codon_start=1
                     /protein_id="CAI19320.1"
                     /db_xref="GI:56203301"
                     /translation="MQESQTKSMFVSRALEKILADKEVKRPQHSQLRRACQVALDEIK
                     AEIEKQRLGTAAPPKANFIEADKYFLPPELACQSKSPRVVSTSLDCLQKLIAYGHITG
                     NAPDSGAPGKRLIDRIVETICSCFQGPQTDEGVQLQIIKALLTAVTSPHIEIHEGTIL
                     QTVRTCYNIYLASKNLINQTTAKATLTQMLNVIFTRMENQVLQEARELEKPIQSKPQS
                     PVIQAAAVSPKFVRLKHSQAQSKPTTPEKTDLTNGEHARSDSGKVSTENGDAPRERGS
                     SLSGTDDGAQEVVKDILEDVVTSAIKEAAEKHGLTEPERVLGELECQECAIPPGVDEN
                     SQTNGIADDRQSLSSADNLESDAQGHQVAARFSHVLQKDAFLVFRSLCKLSMKPLGEG
                     PPDPKSHELRSKVVSLQLLLSVLQNAGPVFRTHEMFINAIKQYLCVALSKNGVSSVPD
                     VFELSLAIFLTLLSNFKMHLKMQIEVPFKEIFLNILETSTSSFEHRWMVIQTLTRICA
                     DAQCVVDIYVNYDCDLNAANIFERLVNDLSKIAQGRSGHELGMTPLQELSLRKKGLEC
                     LVSILKCMVEWSKDLYVNPNHQTSLGQERLTDQEIGDGKGLDMARRCSVTSMESTVSS
                     GTQTTVQDDPEQFEVIKQQKEIIEHGIELFNKKPKRGIQFLQEQGMLGTSVEDIAQFL
                     HQEERLDSTQVGDFLGDSARFNKEVMYAYVDQLDFCEKEFVSALRTFLEGFRLPGEAQ
                     KIDRLMEKFAARYIECNQGQTLFASADTAYVLAYSIIMLTTDLHSPQVKNKMTKEQYI
                     KMNRGINDSKDLPEEYLSSIYEEIEGKKIAMKETKELTIATKSTKQNVASEKQRRLLY
                     NLEMEQMAKTAKALMEAVSHAKAPFTSATHLDHVRPMFKLVWTPLLAAYSIGLQNCDD
                     TEVASLCLEGIRCAIRIACIFGMQLERDAYVQALARFSLLTASSSITEMKQKNIDTIK
                     TLITVAHTDGNYLGNSWHEILKCISQLELAQLIGTGVKTRYLSGSGRREREGSLKGHTL
                     AGEEFMGLGLGNLVSGGVDKRQMASFQESVGETSSQSVVVAVDRIFTGSTRLDGNAIV
                     DFVRWLCAVSMDELASPHHPRMFSLQKIVEISYYNMNRIRLQWSRIWHVIGDHFNKVG
                     CNPNEDVAIFAVDSLRQLSMKFLEKGELANFRFQKDFLRPFEHIMKKNRSPTIRDMAI
                     RCIAQMVNSQAANIRSGWKNIFAVFHQAASDHDGNIVELAFQTTCHIVTTIFQHHFPA
                     AIDSFQDAVKCLSEFACNAAFPDTSMEAIRLIRFCGKYVSERPRVLQEYTSDDMNVAP
                     GDRVWVRGWFPILFELSCIINRCKLDVRTRGLTVMFEIMKSYGHTFEKHWWQDLFRIV
                     FRIFDNMKLPEQLSEKSEWMTTTCNHALYAICDVFTQFYEALNEVLLSDVFAQLQWCV
                     KQDNEQLARSGTNCLENLVISNGEKFSPEVWDETCNCMLDIFKTTIPHVLLTWRPVGM
                     EEDSSEKHLDVDLDRQSLSSIDKNPSERGQSQLSNPTDDSWKGRPYANQKLFASLLIK
                     CVVQLELIQTIDNIVFYPATSKKEDAEHMVAAQQDTLDADIHIETEDQGMYKYMSSQH
                     LFKLLDCLQESHSFSKAFNSNYEQRTVLWRAGFKGKSKPNLLKQETSSLACCLRILFR
                     MYVDENRRDSWEEIQQRLLTVCSEALAYFITVNSESHREAWTSLLLLLLTKTLKINDE
                     KFKAHASMYYPYLCEIMQFDLIPELRAVLRKFFLRIGVVYKIWIPEEPSQVPAALSPV
                     W"
     misc_feature    49711
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /note="Clone_left_end: RP1-155G6"
     misc_feature    69981
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 360bp by BamHI digest data and pUC bridge
                     sizing."
     misc_feature    71108. .71162
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /note="Single clone region. ALU region. pUC subcloned to
                     give average 20x coverage. Assembly confirmed by
                     restriction enzyme digest data."
     misc_feature    71259. .71314
```

```
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-001"
                     /note="Single clone region. ALU region. pUC subcloned to
                     give average 4x coverage. Assembly confirmed by
                     restriction enzyme digest data."
.    gene            join(108600. .108733,108836. .108946)
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-002"
     mRNA            join(108600. .108733,108836. .108946)
                     /gene="ARFGEF2"
                     /locus_tag="RP1-155G6.3-002"
                     /note="match: ESTs: BF833823.1"
     misc_feature    110028
                     /note="Clone_right_end: RP5-1164I10"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 110028;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6165 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6114


RESULT 98
AL590788
LOCUS       AL590788               110042 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-210P2 on chromosome 6, complete
            sequence.
ACCESSION   AL590788
VERSION     AL590788.8  GI:14455976
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110042)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 14, 2001 this sequence version replaced gi:14148913.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-210P2 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .110042
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-210P2"
                     /clone_lib="RPCI-11.1"
     misc_feature    100
                     /note="Clone_right_end: RP3-514B11"
     misc_feature    25329. .25397
                     /note="Sequence from uni-directional dGTP BigDye
                     Terminator reads only."
     misc_feature    110042
                     /note="Clone_right_end: RP11-210P2"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 110042;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63836 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 63887


RESULT 99
AF322449/c
LOCUS       AF322449               111051 bp    DNA     linear   HTG 06-DEC-2000
DEFINITION  Homo sapiens chromosome 17 clone BAC347P24 map 17p13.3, ***
            SEQUENCING IN PROGRESS ***, 26 unordered pieces.
ACCESSION   AF322449
VERSION     AF322449.1  GI:11559854
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 111051)
  AUTHORS   Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
            Gu,J.R.
  TITLE     Gene clone on human chromosome 17p13.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 111051)
  AUTHORS   Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
            Gu,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2000) National Laboratory For Oncogenes & Related
            Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
            200032, P. R. China
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 26 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      484: contig of 484 bp in length
            *                    gap of unknown length
            *      485     1176: contig of 692 bp in length
            *                    gap of unknown length
            *     1177     1450: contig of 274 bp in length
            *                    gap of unknown length
            *     1451     2440: contig of 990 bp in length
            *                    gap of unknown length
            *     2441     3647: contig of 1207 bp in length
            *                    gap of unknown length
            *     3648     4230: contig of 583 bp in length
            *                    gap of unknown length
            *     4231     5202: contig of 972 bp in length
            *                    gap of unknown length
            *     5203     6225: contig of 1023 bp in length
            *                    gap of unknown length
```

```
            *     6226     7535: contig of 1310 bp in length
            *                  gap of unknown length
            *     7536     8924: contig of 1389 bp in length
            *                  gap of unknown length
            *     8925    10775: contig of 1851 bp in length
            *                  gap of unknown length
            *    10776    13242: contig of 2467 bp in length
            *                  gap of unknown length
            *    13243    15999: contig of 2757 bp in length
            *                  gap of unknown length
            *    16000    19279: contig of 3280 bp in length
            *                  gap of unknown length
            *    19280    22810: contig of 3531 bp in length
            *                  gap of unknown length
            *    22811    26950: contig of 4140 bp in length
            *                  gap of unknown length
            *    26951    30765: contig of 3815 bp in length
            *                  gap of unknown length
            *    30766    35709: contig of 4944 bp in length
            *                  gap of unknown length
            *    35710    42248: contig of 6539 bp in length
            *                  gap of unknown length
            *    42249    49778: contig of 7530 bp in length
            *                  gap of unknown length
            *    49779    56400: contig of 6622 bp in length
            *                  gap of unknown length
            *    56401    65978: contig of 9578 bp in length
            *                  gap of unknown length
            *    65979    76311: contig of 10333 bp in length
            *                  gap of unknown length
            *    76312    84457: contig of 8146 bp in length
            *                  gap of unknown length
            *    84458    97903: contig of 13446 bp in length
            *                  gap of unknown length
            *    97904   111051: contig of 13148 bp in length.
FEATURES             Location/Qualifiers
     source          1. .111051
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17p13.3"
                     /clone="BAC347P24"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 111051;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108154 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 108103


RESULT 100
AC114481/c
LOCUS       AC114481               116888 bp    DNA     linear   PRI 29-APR-2002
DEFINITION  Homo sapiens chromosome 3 clone RP13-487A18, complete sequence.
ACCESSION   AC114481
VERSION     AC114481.2  GI:20336790
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 116888)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 116888)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 116888)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
            Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Apr 29, 2002 this sequence version replaced gi:19310300.
            ---------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
            ---------- Project Information
                Center project name: chr-3
                Center clone name: RP13-487A18 (bc0712)
            ---------- Summary Statistics
                Sequencing vector: plasmid; 100% of reads
                Chemistry: Dye-terminator ET; 57% of reads
                Chemistry: Dye-terminator Big Dye; 43% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 116857 bases at least Q40
                Consensus quality: 116886 bases at least Q30
                Consensus quality: 116888 bases at least Q20
                Insert size: 116888; sum-of-contigs
                Quality coverage: 11.8x in Q20 bases; sum-of-contigs
            ----------.
            Overlapping Sequences:
                5':  Mapping in progress
                3':  RP11-151D23 (UWGC:bc0270) AC067763
            --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.

            This sequence was finished as follows unless otherwise noted:
            all regions were either double-stranded or sequenced with an
            alternate chemistry or covered by high quality data (i.e., Phred
            quality >= 30); an attempt was made to resolve all sequencing
            problems, such as compressions and repeats; all regions were
            covered by at least one plasmid subclone or more than one M13
            subclone; and the assembly was confirmed by restriction digest.
            --------------------
            Sequence Validation:
             This sequence has been validated by Multiple Complete Digest
             fingerprinting. Comparison of the experimentally derived digest
             fragments with sequence-predicted fragments is given below.
             The electronically-digested sequence consists of both insert and
             vector, in order to accurately represent the entire circular BAC.
             Small fragments below a variable cutoff (approximately 400-800 bp)
             are not resolved in the fingerprint and hence do not appear
             in the table. There are no significant remaining discrepancies
             between the experimental and predicted values. Uniquely ordered
             fragments are separated by dashed lines.
                   EcoRI                  BglII                  HindIII
            ------------------   ------------------   ------------------
            SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt
            ---------  --------   ---------  --------   ---------  --------

            ---------             ---------             ---------
                16567     16041        8281        8736        7097        7100
            ---------             ---------             ---------
                  885       865        2067        2037         449        <800
```

566	<800	5842	6097	510	<800
10138	10241	726	722	6511	6403
5932	6176	8963	9441	1997	1982
3547	3470	22693	22397	22133	22287
5667	5909	9245	9441	5286	5248
3422	3470	6720	7134	5832	5854
12502	12708	3489	3562	1444	1430
9097	9443	15829	16232	7147	7100
3849	3878	3900	4061	6905	7100
9736	10241	81	<800	2371	2343
3368	3470	12595	12874	4605	4618
128	<800	10029	10085	12762	13005
6610	6926	2887	2925	3492	3534
510	<800	334	<800	5177	5248
15047	16041	2471	2532	3147	3118
18116	18484	9535	10085	3049	3118
				4670	4618
				13645	13005
				7458	7605

```
FEATURES             Location/Qualifiers
     source          1. .116888
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP13-487A18"
                     /clone_lib="null"
     misc_feature    1. .5
                     /note="Single subclone region"
     misc_feature    116885. .116888
                     /note="Single subclone region"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 116888;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26574 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 26523


RESULT 101
AC016590
LOCUS       AC016590               117899 bp    DNA     linear   PRI 22-MAR-2003
DEFINITION  Homo sapiens chromosome 19 clone CTD-3220F14, complete sequence.
ACCESSION   AC016590
VERSION     AC016590.7  GI:28191395
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 117899)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 117899)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 117899)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 117899)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 1, 2003 this sequence version replaced gi:13699590.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.6.
            NOTE: Large tandem repeat 57800-92700. There are possibly more
            repeat copies than are represented in this assembly.
            NOTE: This insert is not the entire sequence of the clone (entire
            sequence is 200kb). It is clipped at the overlaps with AC008806 and
            AC012309. The number of bases overlapped with AC008806 is 7498 bps
            and with AC012309 is 13284 bps.
FEATURES             Location/Qualifiers
     source          1. .117899
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-3220F14"
     misc_feature    57800. .92700
                     /note="NOTE: Large tandem repeat 57800-92700. There are
                     possibly more repeat copies than are represented in this
                     assembly."
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 117899;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8892 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8943


RESULT 102
AC104051/c
LOCUS       AC104051               120723 bp    DNA     linear   PRI 02-MAR-2002
DEFINITION  Homo sapiens chromosome 8, clone CTD-2339F6, complete sequence.
ACCESSION   AC104051
VERSION     AC104051.3  GI:19071664
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 120723)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone CTD-2339F6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 120723)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
```

```
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 120723)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 2, 2002 this sequence version replaced gi:18450132.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L22072
                Center clone name: 2339_F_6
            --------------.
FEATURES             Location/Qualifiers
     source          1. .120723
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="CTD-2339F6"
                     /clone_lib="CITD1 Human BAC"
     repeat_region   35. .813
                     /rpt_family="L1M4c"
     repeat_region   857. .1128
                     /rpt_family="L1M4c"
     repeat_region   1183. .1373
                     /rpt_family="L1ME2"
     repeat_region   1767. .1925
                     /rpt_family="L2"
     repeat_region   2029. .2076
                     /rpt_family="AT_rich"
     repeat_region   2097. .2120
                     /rpt_family="AT_rich"
     repeat_region   complement(3952. .4083)
                     /rpt_family="MER5A"
     repeat_region   complement(4837. .4932)
                     /rpt_family="MIR"
     repeat_region   complement(6143. .6287)
                     /rpt_family="MIR"
     repeat_region   6340. .6709
                     /rpt_family="MSTB"
     repeat_region   complement(6710. .7011)
                     /rpt_family="AluSx"
     repeat_region   7012. .7052
                     /rpt_family="MSTB"
     repeat_region   complement(8754. .8908)
                     /rpt_family="MLT1L"
     repeat_region   complement(9576. .9818)
                     /rpt_family="L1PA8"
     repeat_region   complement(9819. .11131)
                     /rpt_family="L1PA8"
     repeat_region   complement(11132. .11413)
                     /rpt_family="AluSp"
     repeat_region   complement(11414. .15347)
                     /rpt_family="L1PA8"
     repeat_region   15504. .15552
                     /rpt_family="L2"
     repeat_region   complement(16682. .17158)
                     /rpt_family="LTR33"
     repeat_region   17190. .17237
                     /rpt_family="(TG)n"
     repeat_region   17266. .17289
                     /rpt_family="AT_rich"
     repeat_region   complement(19324. .19610)
                     /rpt_family="AluSx"
     repeat_region   20277. .20311
                     /rpt_family="AT_rich"
     repeat_region   21092. .21151
                     /rpt_family="(CA)n"
     repeat_region   22789. .23255
                     /rpt_family="L2"
     repeat_region   complement(23674. .25756)
                     /rpt_family="L1MB7"
     repeat_region   complement(25757. .25934)
                     /rpt_family="L1M4c"
     repeat_region   complement(25935. .26325)
                     /rpt_family="MLT1B"
     repeat_region   complement(26326. .27976)
                     /rpt_family="L1M4c"
     repeat_region   complement(27977. .28340)
                     /rpt_family="MLT1A1"
     repeat_region   complement(28341. .28642)
                     /rpt_family="L1M4c"
     repeat_region   28647. .28700
                     /rpt_family="(TG)n"
     repeat_region   complement(30192. .30366)
                     /rpt_family="AluY"
     repeat_region   complement(30367. .33534)
                     /rpt_family="L1PA3"
     repeat_region   complement(33535. .33656)
                     /rpt_family="AluY"
     repeat_region   33967. .34276
                     /rpt_family="AluSx"
     repeat_region   37011. .37040
                     /rpt_family="AT_rich"
```

```
     repeat_region   complement(37257. .37315)
                     /rpt_family="L2"
     repeat_region   complement(37328. .37491)
                     /rpt_family="MIR"
     repeat_region   37609. .37814
                     /rpt_family="MLT1H1"
     repeat_region   38136. .38168
                     /rpt_family="AT_rich"
     repeat_region   39897. .40222
                     /rpt_family="AluSx"
     repeat_region   40380. .40681
                     /rpt_family="AluSx"
     repeat_region   40700. .40820
                     /rpt_family="MER103"
     unsure          42029. .42156
                     /note="single clone coverage"
     unsure          42157. .42165
                     /note="<30 qual single clone coverage"
     unsure          42166. .42182
                     /note="single clone coverage"
     unsure          42203. .42208
                     /note="<30 qual SNGL region"
     unsure          42234. .42240
                     /note="<30 qual SNGL region"
     repeat_region   44201. .44353
                     /rpt_family="MIR"
     repeat_region   complement(45240. .45375)
                     /rpt_family="MIR3"
     repeat_region   complement(46587. .47079)
                     /rpt_family="LTR40a"
     repeat_region   complement(47685. .47769)
                     /rpt_family="L1MC/D"
     repeat_region   47887. .48226
                     /rpt_family="AluJo"
     repeat_region   49740. .49784
                     /rpt_family="AT_rich"
     repeat_region   complement(51246. .51358)
                     /rpt_family="L2"

  Query Match             1.7%;  Score 52;  DB 8;  Length 120723;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6953 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 6902


RESULT 103
AC025473/c
LOCUS       AC025473              128540 bp    DNA     linear   HTG 05-APR-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11-183N2, WORKING DRAFT SEQUENCE,
            11 unordered pieces.
ACCESSION   AC025473
VERSION     AC025473.2  GI:7417674
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 128540)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 128540)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 5, 2000 this sequence version replaced gi:7211967.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------Summary Statistics
            Consensus quality: 118729 bases at least Q40
            Consensus quality: 125307 bases at least Q30
            Consensus quality: 126460 bases at least Q20
            Estimated insert size: 128540; sum-of-contigs estimation
            Estimated insert size: 202600; agarose-fp estimation
            Quality coverage:  3.58x in Q20 bases; agarose-fp estimation
            Quality coverage:  5.64x in Q20 bases; sum-of-contigs estimation
            -------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2361: contig of 2361 bp in length
            *                   gap of unknown length
            *     2362     4766: contig of 2405 bp in length
            *                   gap of unknown length
            *     4767     7300: contig of 2534 bp in length
            *                   gap of unknown length
            *     7301    10400: contig of 3100 bp in length
            *                   gap of unknown length
            *    10401    14252: contig of 3852 bp in length
            *                   gap of unknown length
            *    14253    20943: contig of 6691 bp in length
            *                   gap of unknown length
            *    20944    26181: contig of 5238 bp in length
            *                   gap of unknown length
            *    26182    32137: contig of 5956 bp in length
            *                   gap of unknown length
            *    32138    47316: contig of 15179 bp in length
            *                   gap of unknown length
            *    47317    71055: contig of 23739 bp in length
            *                   gap of unknown length
            *    71056   128540: contig of 57485 bp in length.
FEATURES             Location/Qualifiers
     source          1. .128540
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-183N2"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 128540;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88327 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 88276


RESULT 104
AL590489/c
LOCUS       AL590489              130146 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-412O4 on chromosome 6q16.2-21
            Contains part of the C6orf210 gene for chromosome 6 open reading
            frame 210, complete sequence.
ACCESSION   AL590489
VERSION     AL590489.11  GI:14625660
KEYWORDS    HTG; C6orf210.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 130146)
```

```
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 6, 2001 this sequence version replaced gi:14575385.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-412O4 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .130146
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIB753O04412"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="q16.2-21"
                     /clone="RP11-412O4"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-412O4"
     gene            join(AL121957.15:80152. .80726,complement(77547. .77681),
                     complement(17378. .17576),complement(7535. .7861))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
     mRNA            join(AL121957.15:80152. .80726,complement(77547. .77681),
                     complement(17378. .17576),complement(7535. .7861))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
                     /product="chromosome 6 open reading frame 210"
                     /note="match: cDNAs: AF086233 BC029491"
     polyA_site      complement(7535)
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
     polyA_site      complement(7537)
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
     polyA_signal    complement(7549. .7554)
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
     CDS             join(AL121957.15:80431. .80726,complement(77547. .77681),
                     complement(17378. .17576),complement(7769. .7861))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-002"
                     /standard_name="OTTHUMP00000065665"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 210"
                     /protein_id="CAH73572.1"
                     /db_xref="GI:55665806"
                     /db_xref="GOA:Q5VU39"
                     /db_xref="UniProt/TrEMBL:Q5VU39"
                     /translation="MNFRQLLLHLPRYLGASGSPRRLWWSPSLDTISSVGSWRGRSSK
                     SPAHWNQVVSEAEKIVGYPTSFMSLRCLLSDELSNIAMQVRKLVGTQHPLLTTARGLV
                     HDSWNSLQLRGLVVLLISKAAGPSSVNTSCQNYDMVSGIYSCQRSLAEITELIHIALL
                     VHRGIVNLNELQSSDGPLKDMQFGNKIAILSGDFLLANACNGLALLQNTKSFSFNGPI
                     AIYQMGDCESAWILSKHPRALS"
     gene            join(AL121957.15:80153. .80726,complement(77547. .77681),
                     complement(17378. .17576),
                     complement(AL591516.9:71178. .71249),
                     complement(AL591516.9:37739. .37912),
                     complement(AL591516.9:36069. .36200),
                     complement(AL591516.9:19452. .19484),
                     complement(AL355586.27:161497. .163717))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-001"
     mRNA            join(AL121957.15:80153. .80726,complement(77547. .77681),
                     complement(17378. .17576),
                     complement(AL591516.9:71178. .71249),
                     complement(AL591516.9:37739. .37912),
                     complement(AL591516.9:36069. .36200),
                     complement(AL591516.9:19452. .19484),
                     complement(AL355586.27:161497. .163717))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-001"
                     /product="chromosome 6 open reading frame 210"
                     /note="match: ESTs: AA007421 AA007488 AA633846 AI033207
                     AI097731 AI492248 AI521803 AI678112 AI765825 AI796241
                     AI827607 AW243064 AW276704 AW408635 AW575740 AW619911
                     BE099861 BE101041 BE221188 BE283879 BE617435 BE818989
                     BF590925 BG292722 BG401578 BI158078 BI465307 BI551760
                     BJ002337 BJ006031 BJ064031 BM017975 BM488969 BM796326
                     BM822264 BM836490 BM974205 H49565 H49821
                     match: cDNAs: AF254956 AF254956.1 AJ420575.1 AK000293
                     AK017329"
     CDS             join(AL121957.15:80431. .80726,complement(77547. .77681),
                     complement(17378. .17576),
                     complement(AL591516.9:71178. .71249),
                     complement(AL591516.9:37739. .37912),
                     complement(AL591516.9:36069. .36200),
                     complement(AL591516.9:19452. .19484),
                     complement(AL355586.27:163559. .163717))
                     /gene="C6orf210"
                     /locus_tag="RP11-59I9.3-001"
                     /standard_name="OTTHUMP00000065664"
                     /note="match: proteins: Q9D3K7 Q9NR58 Q9VP87"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 210"
                     /protein_id="CAH73573.1"
                     /db_xref="GI:55665807"
                     /db_xref="GOA:Q5VU38"
                     /db_xref="UniProt/TrEMBL:Q5VU38"
                     /translation="MNFRQLLLHLPRYLGASGSPRRLWWSPSLDTISSVGSWRGRSSK
                     SPAHWNQVVSEAEKIVGYPTSFMSLRCLLSDELSNIAMQVRKLVGTQHPLLTTARGLV
                     HDSWNSLQLRGLVVLLISKAAGPSSVNTSCQNYDMVSGIYSCQRSLAEITELIHIALL
                     VHRGIVNLNELQSSDGPLKDMQFGNKIAILSGDFLLANACNGLALLQNTKVVELLASA
                     LMDLVQGVYHENSTSKESYITDDIGISTWKEQTFLSHGALLAKSCQAAMELAKHDAEV
                     QNMAFQYGKHMAMSHKINSDVQPFIKEKTSDSMTFNLNSAPVVLHQEFLGRDLWIKQI
                     GEAQEKGRLDYAKLRERIKAGKGVTSAIDLCRYHGNKALEALESFPPSEARSALENIV
                     FAVTRFS"
     misc_feature    72128
                     /note="Clone_right_end: RP11-300D11"
     misc_feature    128147
                     /note="Clone_left_end: RP1-67A8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 130146;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69521 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 69470
```

```
RESULT 105
AC024096
LOCUS       AC024096               130572 bp    DNA     linear   HTG 08-JAN-2003
DEFINITION  Homo sapiens chromosome 3 clone RP11-214A22, WORKING DRAFT
            SEQUENCE, 20 unordered pieces.
ACCESSION   AC024096
VERSION     AC024096.16  GI:20335482
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 130572)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 130572)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 130572)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Apr 28, 2002 this sequence version replaced gi:16117892.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HAGX
                Center clone name: RP11-214A22
            -------------- Summary Statistics
                Sequencing vector: M13;
                Chemistry: Dye-primer Bodipy: 5% of reads
                Chemistry: Dye-terminator Big Dye: 95% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 106061 bases at least Q40
                Consensus quality: 115099 bases at least Q30
                Consensus quality: 121518 bases at least Q20
                Estimated insert size: 143263; sum-of-contigs estimation
                Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2636: contig of 2636 bp in length
            *     2637     2736: gap of unknown length
            *     2737     5887: contig of 3151 bp in length
            *     5888     5987: gap of unknown length
            *     5988     8254: contig of 2267 bp in length
            *     8255     8354: gap of unknown length
            *     8355    10964: contig of 2610 bp in length
            *    10965    11064: gap of unknown length
            *    11065    13703: contig of 2639 bp in length
            *    13704    13803: gap of unknown length
            *    13804    16745: contig of 2942 bp in length
            *    16746    16845: gap of unknown length
            *    16846    20066: contig of 3221 bp in length
            *    20067    20166: gap of unknown length
            *    20167    25220: contig of 5054 bp in length
            *    25221    25320: gap of unknown length
            *    25321    29845: contig of 4525 bp in length
            *    29846    29945: gap of unknown length
            *    29946    35872: contig of 5927 bp in length
            *    35873    35972: gap of unknown length
            *    35973    39682: contig of 3710 bp in length
            *    39683    39782: gap of unknown length
            *    39783    44248: contig of 4466 bp in length
            *    44249    44348: gap of unknown length
            *    44349    50564: contig of 6216 bp in length
            *    50565    50664: gap of unknown length
            *    50665    59327: contig of 8663 bp in length
            *    59328    59427: gap of unknown length
            *    59428    66662: contig of 7235 bp in length
            *    66663    66762: gap of unknown length
            *    66763    74213: contig of 7451 bp in length
            *    74214    74313: gap of unknown length
            *    74314    86510: contig of 12197 bp in length
            *    86511    86610: gap of unknown length
            *    86611   102506: contig of 15896 bp in length
            *   102507   102606: gap of unknown length
            *   102607   115136: contig of 12530 bp in length
            *   115137   115236: gap of unknown length
            *   115237   130572: contig of 15336 bp in length.
FEATURES             Location/Qualifiers
     source          1. .130572
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-214A22"
     gap             2637. .2736
```

```
                 /estimated_length=unknown
     gap         5888. .5987
                 /estimated_length=unknown
     gap         8255. .8354
                 /estimated_length=unknown
     gap         10965. .11064
                 /estimated_length=unknown
     gap         13704. .13803
                 /estimated_length=unknown
     gap         16746. .16845
                 /estimated_length=unknown
     gap         20067. .20166
                 /estimated_length=unknown
     gap         25221. .25320
                 /estimated_length=unknown
     gap         29846. .29945
                 /estimated_length=unknown
     gap         35873. .35972
                 /estimated_length=unknown
     gap         39683. .39782
                 /estimated_length=unknown
     gap         44249. .44348
                 /estimated_length=unknown
     gap         50565. .50664
                 /estimated_length=unknown
     gap         59328. .59427
                 /estimated_length=unknown
     gap         66663. .66762
                 /estimated_length=unknown
     gap         74214. .74313
                 /estimated_length=unknown
     gap         86511. .86610
                 /estimated_length=unknown
     gap         102507. .102606
                 /estimated_length=unknown
     gap         115137. .115236
                 /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 130572;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     113714 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 113765


RESULT 106
AC022766
LOCUS       AC022766              130984 bp    DNA     linear   HTG 01-JUN-2000
DEFINITION  Homo sapiens clone RP11-318E23, WORKING DRAFT SEQUENCE, 31
            unordered pieces.
ACCESSION   AC022766
VERSION     AC022766.3  GI:8138013
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 130984)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-318E23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 130984)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 1, 2000 this sequence version replaced gi:7712161.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5852
                Center clone name: 318_E_23
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 118950 bases at least Q40
                Consensus quality: 124428 bases at least Q30
                Consensus quality: 126573 bases at least Q20
                Insert size: 152000; agarose-fp
                Insert size: 127984; sum-of-contigs
                Quality coverage: 4.6 in Q20 bases; agarose-fp
                Quality coverage: 5.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 31 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1066: contig of 1066 bp in length
            *     1067     1166: gap of 100 bp
            *     1167     2747: contig of 1581 bp in length
            *     2748     2847: gap of 100 bp
            *     2848     4584: contig of 1737 bp in length
            *     4585     4684: gap of 100 bp
            *     4685     6489: contig of 1805 bp in length
            *     6490     6589: gap of 100 bp
            *     6590     8380: contig of 1791 bp in length
            *     8381     8480: gap of 100 bp
            *     8481    10218: contig of 1738 bp in length
            *    10219    10318: gap of 100 bp
            *    10319    12338: contig of 2020 bp in length
            *    12339    12438: gap of 100 bp
            *    12439    14278: contig of 1840 bp in length
            *    14279    14378: gap of 100 bp
            *    14379    16225: contig of 1847 bp in length
            *    16226    16325: gap of 100 bp
            *    16326    18627: contig of 2302 bp in length
            *    18628    18727: gap of 100 bp
            *    18728    21450: contig of 2723 bp in length
            *    21451    21550: gap of 100 bp
            *    21551    25430: contig of 3880 bp in length
            *    25431    25530: gap of 100 bp
            *    25531    28870: contig of 3340 bp in length
            *    28871    28970: gap of 100 bp
            *    28971    32940: contig of 3970 bp in length
            *    32941    33040: gap of 100 bp
```

```
            *    33041    35975: contig of 2935 bp in length
            *    35976    36075: gap of 100 bp
            *    36076    39460: contig of 3385 bp in length
            *    39461    39560: gap of 100 bp
            *    39561    43197: contig of 3637 bp in length
            *    43198    43297: gap of 100 bp
            *    43298    46333: contig of 3036 bp in length
            *    46334    46433: gap of 100 bp
            *    46434    51625: contig of 5192 bp in length
            *    51626    51725: gap of 100 bp
            *    51726    55173: contig of 3448 bp in length
            *    55174    55273: gap of 100 bp
            *    55274    59400: contig of 4127 bp in length
            *    59401    59500: gap of 100 bp
            *    59501    64807: contig of 5307 bp in length
            *    64808    64907: gap of 100 bp
            *    64908    68668: contig of 3761 bp in length
            *    68669    68768: gap of 100 bp
            *    68769    72524: contig of 3756 bp in length
            *    72525    72624: gap of 100 bp
            *    72625    77709: contig of 5085 bp in length
            *    77710    77809: gap of 100 bp
            *    77810    83970: contig of 6161 bp in length
            *    83971    84070: gap of 100 bp
            *    84071    89106: contig of 5036 bp in length
            *    89107    89206: gap of 100 bp
            *    89207    96647: contig of 7441 bp in length
            *    96648    96747: gap of 100 bp
            *    96748   105147: contig of 8400 bp in length
            *   105148   105247: gap of 100 bp
            *   105248   118840: contig of 13593 bp in length
            *   118841   118940: gap of 100 bp
            *   118941   130984: contig of 12044 bp in length.
FEATURES             Location/Qualifiers
     source          1. .130984
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-318E23"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1066
                     /note="assembly_fragment"
     gap             1067. .1166
                     /estimated_length=100
     misc_feature    1167. .2747
                     /note="assembly_fragment"
     gap             2748. .2847
                     /estimated_length=100
     misc_feature    2848. .4584
                     /note="assembly_fragment"
     gap             4585. .4684
                     /estimated_length=100
     misc_feature    4685. .6489
                     /note="assembly_fragment"
     gap             6490. .6589
                     /estimated_length=100
     misc_feature    6590. .8380
                     /note="assembly_fragment"
     gap             8381. .8480
                     /estimated_length=100
     misc_feature    8481. .10218
                     /note="assembly_fragment"
     gap             10219. .10318
                     /estimated_length=100
     misc_feature    10319. .12338
                     /note="assembly_fragment"
     gap             12339. .12438
                     /estimated_length=100
     misc_feature    12439. .14278
                     /note="assembly_fragment"
     gap             14279. .14378
                     /estimated_length=100
     misc_feature    14379. .16225
                     /note="assembly_fragment"
     gap             16226. .16325
                     /estimated_length=100
     misc_feature    16326. .18627
                     /note="assembly_fragment"
     gap             18628. .18727
                     /estimated_length=100
     misc_feature    18728. .21450
                     /note="assembly_fragment"
     gap             21451. .21550
                     /estimated_length=100
     misc_feature    21551. .25430
                     /note="assembly_fragment"
     gap             25431. .25530
                     /estimated_length=100
     misc_feature    25531. .28870
                     /note="assembly_fragment"
     gap             28871. .28970
                     /estimated_length=100
     misc_feature    28971. .32940
                     /note="assembly_fragment"
     gap             32941. .33040
                     /estimated_length=100
     misc_feature    33041. .35975
                     /note="assembly_fragment"
     gap             35976. .36075
                     /estimated_length=100
     misc_feature    36076. .39460
                     /note="assembly_fragment"
     gap             39461. .39560
                     /estimated_length=100

  Query Match             1.7%;  Score 52;  DB 14;  Length 130984;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18564 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 18615


RESULT 107
AC100793/c
LOCUS       AC100793              134465 bp    DNA     linear   PRI 15-JAN-2003
DEFINITION  Homo sapiens chromosome 17, clone CTD-3193K9, complete sequence.
ACCESSION   AC100793
VERSION     AC100793.8  GI:27753789
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 134465)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone CTD-3193K9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 134465)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
```

```
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 134465)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 134465)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 15, 2003 this sequence version replaced gi:24415443.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20710
                Center clone name: 3193_K_9
            --------------.
FEATURES             Location/Qualifiers
     source          1. .134465
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="CTD-3193K9"
                     /clone_lib="CITD2 Human BAC"
     repeat_region   1. .88
                     /rpt_family="AluJo/FRAM"
     repeat_region   94. .389
                     /rpt_family="AluSx"
     repeat_region   complement(479. .551)
                     /rpt_family="Tigger5"
     repeat_region   575. .694
                     /rpt_family="MER21B"
     repeat_region   1296. .1431
                     /rpt_family="AluJb"
     repeat_region   1432. .1728
                     /rpt_family="AluSx"
     repeat_region   1731. .2062
                     /rpt_family="AluJb"
     repeat_region   2400. .2547
                     /rpt_family="L2"
     repeat_region   2613. .2661
                     /rpt_family="(TA)n"
     repeat_region   complement(2670. .2789)
                     /rpt_family="FLAM_C"
     repeat_region   complement(2795. .3122)
                     /rpt_family="AluY"
     repeat_region   6180. .6232
                     /rpt_family="L2"
     repeat_region   complement(6233. .6530)
                     /rpt_family="AluY"
     repeat_region   6531. .6585
                     /rpt_family="L2"
     repeat_region   complement(6586. .6887)
                     /rpt_family="AluSc"
     repeat_region   6888. .7159
                     /rpt_family="L2"
     repeat_region   7875. .7912
                     /rpt_family="GC_rich"
     repeat_region   7957. .8044
                     /rpt_family="(CGGGG)n"
     repeat_region   9590. .9691
                     /rpt_family="MER91A"
     repeat_region   10084. .10105
                     /rpt_family="GC_rich"
     repeat_region   10488. .10589
                     /rpt_family="MIR"
     repeat_region   12437. .12588
                     /rpt_family="MIR3"
     repeat_region   12868. .12900
                     /rpt_family="(CCCCCT)n"
     repeat_region   13333. .13370
                     /rpt_family="GC_rich"
     repeat_region   14420. .14475
                     /rpt_family="GA-rich"
     repeat_region   14721. .14744
                     /rpt_family="GC_rich"
     repeat_region   14801. .14840
                     /rpt_family="GC_rich"
     repeat_region   complement(15362. .15663)
                     /rpt_family="AluJo"
     repeat_region   15881. .15924
                     /rpt_family="GC_rich"
     repeat_region   16110. .16162
                     /rpt_family="(CGGGG)n"
     repeat_region   complement(16380. .16658)
                     /rpt_family="L3b"
     repeat_region   complement(16694. .16909)
                     /rpt_family="L3"
     repeat_region   17073. .17108
                     /rpt_family="GC_rich"
     repeat_region   17379. .17409
                     /rpt_family="GC_rich"
     repeat_region   18578. .18710
                     /rpt_family="MIR"
```

```
     repeat_region   complement(18775. .19088)
                     /rpt_family="AluSx"
     repeat_region   19780. .19886
                     /rpt_family="GA-rich"
     repeat_region   19935. .20031
                     /rpt_family="MIR"
     repeat_region   20334. .20445
                     /rpt_family="GA-rich"
     repeat_region   20855. .20960
                     /rpt_family="(CTGGGG)n"
     repeat_region   20906. .21085
                     /rpt_family="(TTGGGG)n"
     repeat_region   21037. .21216
                     /rpt_family="(CTGGGG)n"
     repeat_region   21166. .21309
                     /rpt_family="(TTGGGG)n"
     repeat_region   complement(22213. .22416)
                     /rpt_family="L2"
     repeat_region   22417. .22720
                     /rpt_family="AluSg"
     repeat_region   23332. .23457
                     /rpt_family="MIR"
     repeat_region   complement(23477. .23601)
                     /rpt_family="MIR"

  Query Match              1.7%;  Score 52;  DB 8;  Length 134465;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45103 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 45052


RESULT 108
AC069511
LOCUS       AC069511               135805 bp    DNA     linear   HTG 25-JUL-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-224I22, WORKING DRAFT
            SEQUENCE, 26 unordered pieces.
ACCESSION   AC069511
VERSION     AC069511.4  GI:9438659
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 135805)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 135805)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 25, 2000 this sequence version replaced gi:8468798.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HBER
                Center clone name: RP11-224I22
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 113598 bases at least Q40
                Consensus quality: 123530 bases at least Q30
                Consensus quality: 129169 bases at least Q20
                Estimated insert size: 126394; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 3.1x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 26 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    22050: contig of 22050 bp in length
            *    22051    22150: gap of unknown length
            *    22151    34840: contig of 12690 bp in length
            *    34841    34940: gap of unknown length
            *    34941    46514: contig of 11574 bp in length
            *    46515    46614: gap of unknown length
            *    46615    61066: contig of 14452 bp in length
            *    61067    61166: gap of unknown length
            *    61167    68849: contig of 7683 bp in length
            *    68850    68949: gap of unknown length
            *    68950    76117: contig of 7168 bp in length
            *    76118    76217: gap of unknown length
            *    76218    83757: contig of 7540 bp in length
            *    83758    83857: gap of unknown length
            *    83858    89561: contig of 5704 bp in length
            *    89562    89661: gap of unknown length
            *    89662    96097: contig of 6436 bp in length
            *    96098    96197: gap of unknown length
            *    96198   102149: contig of 5952 bp in length
            *   102150   102249: gap of unknown length
            *   102250   106020: contig of 3771 bp in length
            *   106021   106120: gap of unknown length
            *   106121   110133: contig of 4013 bp in length
            *   110134   110233: gap of unknown length
            *   110234   113954: contig of 3721 bp in length
            *   113955   114054: gap of unknown length
            *   114055   116296: contig of 2242 bp in length
            *   116297   116396: gap of unknown length
            *   116397   118231: contig of 1835 bp in length
            *   118232   118331: gap of unknown length
            *   118332   119935: contig of 1604 bp in length
            *   119936   120035: gap of unknown length
            *   120036   121763: contig of 1728 bp in length
            *   121764   121863: gap of unknown length
            *   121864   123764: contig of 1901 bp in length
            *   123765   123864: gap of unknown length
            *   123865   126122: contig of 2258 bp in length
            *   126123   126222: gap of unknown length
            *   126223   127558: contig of 1336 bp in length
            *   127559   127658: gap of unknown length
            *   127659   128808: contig of 1150 bp in length
            *   128809   128908: gap of unknown length
            *   128909   130088: contig of 1180 bp in length
```

```
            *   130089   130188: gap of unknown length
            *   130189   131521: contig of 1333 bp in length
            *   131522   131621: gap of unknown length
            *   131622   133147: contig of 1526 bp in length
            *   133148   133247: gap of unknown length
            *   133248   134356: contig of 1109 bp in length
            *   134357   134456: gap of unknown length
            *   134457   135805: contig of 1349 bp in length.
FEATURES             Location/Qualifiers
     source          1. .135805
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-224I22"
     gap             22051. .22150
                     /estimated_length=unknown
     gap             34841. .34940
                     /estimated_length=unknown
     gap             46515. .46614
                     /estimated_length=unknown
     gap             61067. .61166
                     /estimated_length=unknown
     gap             68850. .68949
                     /estimated_length=unknown
     gap             76118. .76217
                     /estimated_length=unknown
     gap             83758. .83857
                     /estimated_length=unknown
     gap             89562. .89661
                     /estimated_length=unknown
     gap             96098. .96197
                     /estimated_length=unknown
     gap             102150. .102249
                     /estimated_length=unknown
     gap             106021. .106120
                     /estimated_length=unknown
     gap             110134. .110233
                     /estimated_length=unknown
     gap             113955. .114054
                     /estimated_length=unknown
     gap             116297. .116396
                     /estimated_length=unknown
     gap             118232. .118331
                     /estimated_length=unknown
     gap             119936. .120035
                     /estimated_length=unknown
     gap             121764. .121863
                     /estimated_length=unknown
     gap             123765. .123864
                     /estimated_length=unknown
     gap             126123. .126222
                     /estimated_length=unknown
     gap             127559. .127658
                     /estimated_length=unknown
     gap             128809. .128908
                     /estimated_length=unknown
     gap             130089. .130188
                     /estimated_length=unknown
     gap             131522. .131621
                     /estimated_length=unknown
     gap             133148. .133247
                     /estimated_length=unknown
     gap             134357. .134456
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 135805;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      85075 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85126


RESULT 109
AC006006/c
LOCUS       AC006006              136385 bp    DNA     linear   PRI 27-JAN-2004
DEFINITION  Homo sapiens PAC clone RP4-813F11 from 7, complete sequence.
ACCESSION   AC006006
VERSION     AC006006.2  GI:4753279
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 136385)
  AUTHORS   Hillier,L.W., Fulton,R.S., Fulton,L.A., Graves,T.A., Pepin,K.H.,
            Wagner-McPherson,C., Layman,D., Maas,J., Jaeger,S., Walker,R.,
            Wylie,K., Sekhon,M., Becker,M.C., O'Laughlin,M.D., Schaller,M.E.,
            Fewell,G.A., Delehaunty,K.D., Miner,T.L., Nash,W.E., Cordes,M.,
            Du,H., Sun,H., Edwards,J., Bradshaw-Cordum,H., Ali,J., Andrews,S.,
            Isak,A., Vanbrunt,A., Nguyen,C., Du,F., Lamar,B., Courtney,L.,
            Kalicki,J., Ozersky,P., Bielicki,L., Scott,K., Holmes,A.,
            Harkins,R., Harris,A., Strong,C.M., Hou,S., Tomlinson,C.,
            Dauphin-Kohlberg,S., Kozlowicz-Reilly,A., Leonard,S., Rohlfing,T.,
            Rock,S.M., Tin-Wollam,A.M., Abbott,A., Minx,P., Maupin,R.,
            Strowmatt,C., Latreille,P., Miller,N., Johnson,D., Murray,J.,
            Woessner,J.P., Wendl,M.C., Yang,S.P., Schultz,B.R., Wallis,J.W.,
            Spieth,J., Bieri,T.A., Nelson,J.O., Berkowicz,N., Wohldmann,P.E.,
            Cook,L.L., Hickenbotham,M.T., Eldred,J., Williams,D., Bedell,J.A.,
            Mardis,E.R., Clifton,S.W., Chissoe,S.L., Marra,M.A., Raymond,C.,
            Haugen,E., Gillett,W., Zhou,Y., James,R., Phelps,K., Iadanoto,S.,
            Bubb,K., Simms,E., Levy,R., Clendenning,J., Kaul,R., Kent,W.J.,
            Furey,T.S., Baertsch,R.A., Brent,M.R., Keibler,E., Flicek,P.,
            Bork,P., Suyama,M., Bailey,J.A., Portnoy,M.E., Torrents,D.,
            Chinwalla,A.T., Gish,W.R., Eddy,S.R., McPherson,J.D., Olson,M.V.,
            Eichler,E.E., Green,E.D., Waterston,R.H. and Wilson,R.K.
  TITLE     The DNA sequence of human chromosome 7
  JOURNAL   Nature 424 (6945), 157-164 (2003)
   PUBMED   12853948
REFERENCE   2  (bases 1 to 136385)
  AUTHORS   Drone,K., Le T,P. and Wohldmann,P.
  TITLE     The sequence of Homo sapiens PAC clone RP4-813F11
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 136385)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 136385)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 136385)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 136385)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 136385)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2004) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On May 5, 1999 this sequence version replaced gi:3907515.
            -------------- Genome Center
```

```
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: sapiens@watson.wustl.edu
        ------------- Summary Statistics
            Center project name: H_DJ0813F11
        -------------.

        NOTICE:  This sequence may not represent the entire insert of this
        clone.  It may be shorter because we only sequence overlapping
        clone sections once, or longer because we provide a small overlap
        between neighboring data submissions.

        This sequence was finished as follows unless otherwise noted:
        all regions were double stranded, sequenced with an alternate
        chemistry, or covered by high quality data (i.e., phred quality >=
        30); an attempt was made to resolve all sequencing problems, such
        as compressions and repeats; all regions were covered by sequence
        from more than one subclone; and the assembly was confirmed by
        restriction digest.

        MAPPING INFORMATION:
        The sequence of this clone was established as part of a mapping and
        sequencing collaboration between the NHGRI Chromosome 7 Mapping
        Project (Eric D. Green, Director), John D. McPherson in the
        Department of Genetics (Washington University), and the Washington
        University Genome Sequencing Center. For additional information
        about the map position of this sequence, see
        http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
        mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

        SOURCE INFORMATION:
        This clone was derived from human PAC library RPCI-4, prepared by
        Pieter de Jong and coworkers at http://www.chori.org using the
        method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
        The library is from one male donor.
        The clone may be obtained either from Genome Systems, Inc.
        (http://www.genomesystems.com) or Research Genetics, Inc.
        (http://www.resgen.com); or from Pieter de Jong.
        VECTOR:  pCYPAC2

        NEIGHBORING SEQUENCE INFORMATION:
        The clone sequenced to the left is RP5-839B19, 200 bp overlap the
        clone sequenced to the right is RP5-1007M8.  Actual start of this
        clone is at base position 197 of RP4-813F11 actual end is at 136385
        of RP4-813F11.
FEATURES             Location/Qualifiers
     source          1. .136385
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP4-813F11"
                     /clone_lib="RPCI-4"
     repeat_region   1. .91
                     /rpt_family="Alu"
     repeat_region   103. .458
                     /rpt_family="MER2_type"
     repeat_region   452. .538
                     /rpt_family="MER2_type"
     repeat_region   561. .664
                     /rpt_family="Alu"
     repeat_region   665. .1267
                     /rpt_family="MER2_type"
     repeat_region   1342. .1376
                     /rpt_family="Mariner"
     repeat_region   1494. .1799
                     /rpt_family="Alu"
     repeat_region   1833. .2172
                     /rpt_family="Alu"
     repeat_region   2948. .3060
                     /rpt_family="MIR"
     snRNA           3194. .3300
                     /product="snRNA-U6-related"
     repeat_region   3194. .3300
                     /rpt_family="U6"
     repeat_region   3301. .3345
                     /rpt_family="AT_rich"
     repeat_region   3552. .3598
                     /rpt_family="(TG)n"
     repeat_region   3611. .3870
                     /rpt_family="Alu"
     repeat_region   3871. .4096
                     /rpt_family="MER1_type"
     repeat_region   4097. .4424
                     /rpt_family="Alu"
     repeat_region   4425. .4657
                     /rpt_family="MER1_type"
     repeat_region   4660. .4841
                     /rpt_family="L2"
     repeat_region   4921. .5043
                     /rpt_family="L2"
     repeat_region   5310. .5872
                     /rpt_family="L2"
     repeat_region   6007. .6151
                     /rpt_family="L2"
     repeat_region   6843. .7002
                     /rpt_family="Alu"
     repeat_region   7003. .7294
                     /rpt_family="Alu"
     repeat_region   7307. .7672
                     /rpt_family="MER1_type"
     repeat_region   7680. .7971
                     /rpt_family="Alu"
     repeat_region   7981. .8122
                     /rpt_family="MER1_type"
     repeat_region   8154. .8213
                     /rpt_family="MER2_type"
     repeat_region   8214. .8274
                     /rpt_family="MER1_type"
     repeat_region   8333. .8454
                     /rpt_family="MER1_type"
     repeat_region   8997. .9108
                     /rpt_family="MIR"
     repeat_region   9144. .9340
                     /rpt_family="MIR"
     repeat_region   9489. .9600
                     /rpt_family="L2"
     repeat_region   9589. .9732
                     /rpt_family="CR1"
     repeat_region   9735. .9898
                     /rpt_family="MER1_type"
     repeat_region   10317. .10376
                     /rpt_family="L2"
     repeat_region   10406. .10713
                     /rpt_family="MER2_type"
     repeat_region   10720. .10879
                     /rpt_family="MER1_type"
     repeat_region   11049. .11096
                     /rpt_family="MER103"
     repeat_region   11097. .11393

  Query Match             1.7%;  Score 52;  DB 8;  Length 136385;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114432 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 114381


RESULT 110
AC010367/c
LOCUS       AC010367              137845 bp    DNA     linear   HTG 20-APR-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2044L23, WORKING DRAFT
```

```
            SEQUENCE, 17 unordered pieces.
ACCESSION   AC010367
VERSION     AC010367.4  GI:13699509
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 137845)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137845)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 20, 2001 this sequence version replaced gi:7710742.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 644732
            Center clone name: CITB-H1_2044L23
            -------
            Summary Statistics
            Consensus quality: 122756 bases at least Q40
            Consensus quality: 127174 bases at least Q30
            Consensus quality: 129642 bases at least Q20
            Estimated insert size: 121000; pulse field gel estimation
            Estimated insert size: 136245; sum-of-contigs estimation
            Quality coverage:  6.63 in Q20 bases; pulse field gel estimation
            Quality coverage:  5.89 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1041: contig of 1041 bp in length
            *     1042     1141: gap of unknown length
            *     1142     2183: contig of 1042 bp in length
            *     2184     2283: gap of unknown length
            *     2284     3529: contig of 1246 bp in length
            *     3530     3629: gap of unknown length
            *     3630     4897: contig of 1268 bp in length
            *     4898     4997: gap of unknown length
            *     4998     6540: contig of 1543 bp in length
            *     6541     6640: gap of unknown length
            *     6641     7735: contig of 1095 bp in length
            *     7736     7835: gap of unknown length
            *     7836     8852: contig of 1017 bp in length
            *     8853     8952: gap of unknown length
            *     8953    10065: contig of 1113 bp in length
            *    10066    10165: gap of unknown length
            *    10166    11335: contig of 1170 bp in length
            *    11336    11435: gap of unknown length
            *    11436    12444: contig of 1009 bp in length
            *    12445    12544: gap of unknown length
            *    12545    13852: contig of 1308 bp in length
            *    13853    13952: gap of unknown length
            *    13953    15056: contig of 1104 bp in length
            *    15057    15156: gap of unknown length
            *    15157    23003: contig of 7847 bp in length
            *    23004    23103: gap of unknown length
            *    23104    41386: contig of 18283 bp in length
            *    41387    41486: gap of unknown length
            *    41487    68942: contig of 27456 bp in length
            *    68943    69042: gap of unknown length
            *    69043    98301: contig of 29259 bp in length
            *    98302    98401: gap of unknown length
            *    98402   137845: contig of 39444 bp in length.
FEATURES             Location/Qualifiers
     source          1. .137845
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2044L23"
                     /clone_lib="CalTech human BAC library D"
     gap             1042. .1141
                     /estimated_length=unknown
     gap             2184. .2283
                     /estimated_length=unknown
     gap             3530. .3629
                     /estimated_length=unknown
     gap             4898. .4997
                     /estimated_length=unknown
     gap             6541. .6640
                     /estimated_length=unknown
     gap             7736. .7835
                     /estimated_length=unknown
     gap             8853. .8952
                     /estimated_length=unknown
     gap             10066. .10165
                     /estimated_length=unknown
     gap             11336. .11435
                     /estimated_length=unknown
     gap             12445. .12544
                     /estimated_length=unknown
     gap             13853. .13952
                     /estimated_length=unknown
     gap             15057. .15156
                     /estimated_length=unknown
     gap             23004. .23103
                     /estimated_length=unknown
     gap             41387. .41486
                     /estimated_length=unknown
     gap             68943. .69042
                     /estimated_length=unknown
     gap             98302. .98401
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 137845;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33124 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 33073


RESULT 111
AL954211
LOCUS       AL954211              138573 bp    DNA     linear   PRI 16-OCT-2003
DEFINITION  Pan troglodytes chromosome 22 clone PTB-034G05 map 22q22.11,
            complete sequence.
ACCESSION   AL954211
VERSION     AL954211.1  GI:37605794
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 138573)
  AUTHORS   .
  CONSRTM   The Chimpanzee Chromosome 22 Sequencing Consortium
  TITLE     DNA sequence of chimpanzee chromosome 22 and its evolutionary
            implications
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 138573)
  AUTHORS   Kube,M., Sudbrak,R., Borzym,K., Mueller,I., Klages,S., Burgtorf,C.,
            Heitmann,K., Gimmel,V., Beck,A., Ben Kahla,A., Lehrach,H.,
            Yaspo,M.L. and Reinhardt,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2002) Max-Planck-Institute for Molecular
            Genetics, Department of Vertebrate Genomics, Ihnestrasse 73,
            Berlin, 14195 Germany
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
             *Chinese National Human Genome Center at Shanghai,
              Shanghai, China;
             *GBF, Dept. of Genome Analysis, Braunschweig, Germany;
             *Institute of Molecular Biotechnology, Jena, Germany;
             *KRIBB Genome Research Center, Daejeon, Korea;
             *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
             *National Institute of Genetics, Mishima, Japan;
             *National Yang Ming University Genome Research Center, Taipei,
              Taiwan;
             *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
                 Center: Max-Planck-Institute for Molecular Genetics
                 Center code: MPIMG
            -------------- Project Information
                 Center clone name: PTB-034G05
            -------------- Summary Statistics
                 Sequencing vector: pUC18; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990329
                 Consensus quality: 138557 bases at least Q40
                 Consensus quality: 138569 bases at least Q30
                 Consensus quality: 138573 bases at least Q20
                 Quality coverage: 13.72
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. --------------
            Neighboring sequence information:
            This clone is overlapped by
                 RP43-012J05
                 RP43-043L10
            --------------
            SOURCE INFORMATION:
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee. Clones may be obtained from Asao Fujiyama and
            co-workers (http://www.gsc.riken.go.jp).
            VECTOR:  pKS145.
FEATURES                 Location/Qualifiers
     source              1. .138573
                         /organism="Pan troglodytes"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:9598"
                         /chromosome="22"
                         /map="22q22.11"
                         /clone="PTB-034G05"
     source              <1. .12748
                         /organism="Pan troglodytes"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:9598"
                         /clone="RP43-012J05"
                         /note="overlapping clone"
     source              96436. .>138573
                         /organism="Pan troglodytes"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:9598"
                         /clone="RP43-043L10"
                         /note="overlapping clone"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 138573;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112046 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 112097


RESULT 112
AC073542
LOCUS       AC073542              141633 bp    DNA     linear   PRI 01-MAY-2001
DEFINITION  Homo sapiens chromosome 5 clone RP11-221L16, complete sequence.
ACCESSION   AC073542
VERSION     AC073542.4  GI:13876495
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 141633)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 141633)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 141633)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 141633)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 28, 2001 this sequence version replaced gi:13677059.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .141633
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-221L16"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 141633;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105940 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105991


RESULT 113
AL357552/c
LOCUS       AL357552              143969 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP4-633I8 on chromosome 1p36.21-36.32
            Contains the 5' end of the RERE gene for arginine-glutamic acid
            dipeptide (RE) repeats (FLJ38775), two ribosomal protein L7 (RPL7)
            pseudogenes, a ribosomal protein L27 (RPL27) pseudogene, a novel
```

```
            gene and a CpG island, complete sequence.
ACCESSION   AL357552
VERSION     AL357552.25  GI:27652622
KEYWORDS    HTG; FLJ38775; RERE; RPL27; RPL7.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 143969)
  AUTHORS   Davidson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 11, 2003 this sequence version replaced gi:27498402.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP4-633I8 is from the library RPCI-4 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .143969
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704I08633"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p36.21-36.32"
                     /clone="RP4-633I8"
                     /clone_lib="RPCI-4"
     misc_feature    1
                     /note="Clone_left_end: RP4-633I8"
     misc_feature    8344
                     /note="Clone_right_end: RP11-3J23"
     gene            join(complement(16843. .17007),
                     complement(AL357713.11:73394. .73642))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-009"
     mRNA            join(complement(16843. .17007),
                     complement(AL357713.11:73394. .73642))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-009"
                     /product="arginine-glutamic acid dipeptide (RE) repeats"
                     /note="match: ESTs: Em:AI524637.1 Em:AW444820.1
                     Em:BE675021.1 Em:BF512881.1"
     gene            64218. .64875
                     /locus_tag="RP4-633I8.1-001"
                     /pseudo
     CDS             64218. .64875
                     /locus_tag="RP4-633I8.1-001"
                     /note="match: proteins: Sw:P05426 Sw:P14148 Sw:P18124
                     Tr:AAH25909 Tr:BAC34366 Tr:BAC40262 Tr:O95036 Tr:Q15289
                     Tr:Q80UR0 Tr:Q90YW3 Tr:Q9DG69"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L7 (RPL7) pseudogene"
     gene            join(complement(67219. .67626),
                     complement(AL357713.11:73174. .73642),
                     complement(AL357713.11:41511. .41581),
                     complement(AL357713.11:31762. .31887),
                     complement(AL096855.48:152922. .153027),
                     complement(AL096855.48:151979. .152075),
                     complement(AL096855.48:136718. .136822),
                     complement(AL096855.48:104131. .104179),
                     complement(AL096855.48:92910. .93034),
                     complement(AL096855.48:90568. .90667),
                     complement(AL096855.48:60475. .61528))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-010"
     mRNA            join(complement(67219. .67626),
                     complement(AL357713.11:73174. .73642),
                     complement(AL357713.11:41511. .41581),
                     complement(AL357713.11:31762. .31887),
                     complement(AL096855.48:152922. .153027),
                     complement(AL096855.48:151979. .152075),
                     complement(AL096855.48:136718. .136822),
                     complement(AL096855.48:104131. .104179),
                     complement(AL096855.48:92910. .93034),
                     complement(AL096855.48:90568. .90667),
                     complement(AL096855.48:60475. .61528))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-010"
                     /product="arginine-glutamic acid dipeptide (RE) repeats"
                     /note="match: cDNAs: Em:AK096094.1"
     misc_feature    87829. .87916
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    94628
                     /note="Clone_left_end: RP11-185B14"
     misc_feature    96733. .96734
                     /note="Sequence from overlapping clone RP11-185B14
                     (AL354930). Assembly confirmed by restriction digest."
     gene            complement(99991. .100723)
                     /locus_tag="RP4-633I8.2-001"
                     /pseudo
     CDS             complement(99991. .100723)
                     /locus_tag="RP4-633I8.2-001"
                     /note="match: proteins: Sw:P05426 Sw:P14148 Sw:P18124
                     Tr:BAC40262 Tr:CAE02125 Tr:Q9SCF9"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L7 (RPL7) pseudogene\n\"
     gene            join(complement(130948. .131253),
                     complement(106192. .106376),
                     complement(AL357713.11:73174. .73642),
                     complement(AL357713.11:41511. .41581),
                     complement(AL357713.11:31762. .31887),
                     complement(AL096855.48:152922. .153027),
                     complement(AL096855.48:151979. .152075),
                     complement(AL096855.48:136718. .136822),
                     complement(AL096855.48:104131. .104179),
                     complement(AL096855.48:92910. .93034),
                     complement(AL096855.48:90568. .90667),
                     complement(AL096855.48:61430. .61528),
                     complement(AL096855.48:18232. .18312),
                     complement(AL356072.21:81648. .81810),
                     complement(AL356072.21:80582. .80674),
                     complement(AL356072.21:79892. .80091),
                     complement(AL356072.21:78519. .78680),
                     complement(AL356072.21:77599. .77712),
                     complement(AL356072.21:75948. .77326),
                     complement(AL356072.21:75600. .75822),
                     complement(AL356072.21:74032. .74752),
                     complement(AL356072.21:71936. .72082),
```

```
                     complement(AL356072.21:71255. .71435),
                     complement(AL356072.21:68233. .70956))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-001"
     mRNA            join(complement(130948. .131253),
                     complement(106192. .106376),
                     complement(AL357713.11:73174. .73642),
                     complement(AL357713.11:41511. .41581),
                     complement(AL357713.11:31762. .31887),
                     complement(AL096855.48:152922. .153027),
                     complement(AL096855.48:151979. .152075),
                     complement(AL096855.48:136718. .136822),
                     complement(AL096855.48:104131. .104179),
                     complement(AL096855.48:92910. .93034),
                     complement(AL096855.48:90568. .90667),
                     complement(AL096855.48:61430. .61528),
                     complement(AL096855.48:18232. .18312),
                     complement(AL356072.21:81648. .81810),
                     complement(AL356072.21:80582. .80674),
                     complement(AL356072.21:79892. .80091),
                     complement(AL356072.21:78519. .78680),
                     complement(AL356072.21:77599. .77712),
                     complement(AL356072.21:75948. .77326),
                     complement(AL356072.21:75600. .75822),
                     complement(AL356072.21:74032. .74752),
                     complement(AL356072.21:71936. .72082),
                     complement(AL356072.21:71255. .71435),
                     complement(AL356072.21:68233. .70956))
                     /gene="RERE"
                     /locus_tag="RP11-141M15.2-001"
                     /product="arginine-glutamic acid dipeptide (RE) repeats"
                     /note="match: cDNAs: Em:AB036737.1"
     gene            join(112263. .112479,113097. .113294)
                     /locus_tag="RP4-633I8.3-001"
                     /pseudo
     CDS             join(112263. .112479,113097. .113294)
                     /locus_tag="RP4-633I8.3-001"
                     /note="match: proteins: Sw:P08526 Tr:AAH24366 Tr:BAB28321
                     Tr:Q8WTY3 Tr:Q90YU1 Tr:Q9DFL2"
                     /pseudo
                     /codon_start=1
                     /product="ribosomal protein L27 (RPL27) pseudogene"
     misc_feature    115538. .115554
                     /note="Sequence from overlapping clone RP11-185B14
                     (AL354930). Assembly confirmed by restriction digest."
     gene            join(119648. .119726,120566. .120839)
                     /locus_tag="RP4-633I8.4-001"
     mRNA            join(119648. .119726,120566. .120839)
                     /locus_tag="RP4-633I8.4-001"

  Query Match              1.7%;  Score 52;  DB 8;  Length 143969;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50064 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 50013


RESULT 114
AC027146
LOCUS       AC027146               144542 bp    DNA     linear   HTG 28-AUG-2002
DEFINITION  Homo sapiens chromosome 17 clone RP11-123N2 map 17, WORKING DRAFT
            SEQUENCE, 28 unordered pieces.
ACCESSION   AC027146
VERSION     AC027146.1  GI:7331426
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 144542)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-123N2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 144542)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 144542)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5016
                Center clone name: 123_N_2
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 128149 bases at least Q40
                Consensus quality: 136286 bases at least Q30
                Consensus quality: 139579 bases at least Q20
                Insert size: 168000; agarose-fp
                Insert size: 141842; sum-of-contigs
```

```
            Quality coverage: 3.5 in Q20 bases; agarose-fp
            Quality coverage: 4.2 in Q20 bases; sum-of-contigs
           --------------.
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 28 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *        1     1042: contig of 1042 bp in length
           *     1043     1142: gap of 100 bp
           *     1143     2215: contig of 1073 bp in length
           *     2216     2315: gap of 100 bp
           *     2316     3520: contig of 1205 bp in length
           *     3521     3620: gap of 100 bp
           *     3621     6235: contig of 2615 bp in length
           *     6236     6335: gap of 100 bp
           *     6336     8261: contig of 1926 bp in length
           *     8262     8361: gap of 100 bp
           *     8362    10203: contig of 1842 bp in length
           *    10204    10303: gap of 100 bp
           *    10304    11741: contig of 1438 bp in length
           *    11742    11841: gap of 100 bp
           *    11842    13740: contig of 1899 bp in length
           *    13741    13840: gap of 100 bp
           *    13841    16452: contig of 2612 bp in length
           *    16453    16552: gap of 100 bp
           *    16553    19320: contig of 2768 bp in length
           *    19321    19420: gap of 100 bp
           *    19421    21340: contig of 1920 bp in length
           *    21341    21440: gap of 100 bp
           *    21441    23790: contig of 2350 bp in length
           *    23791    23890: gap of 100 bp
           *    23891    25590: contig of 1700 bp in length
           *    25591    25690: gap of 100 bp
           *    25691    28404: contig of 2714 bp in length
           *    28405    28504: gap of 100 bp
           *    28505    30466: contig of 1962 bp in length
           *    30467    30566: gap of 100 bp
           *    30567    34336: contig of 3770 bp in length
           *    34337    34436: gap of 100 bp
           *    34437    37626: contig of 3190 bp in length
           *    37627    37726: gap of 100 bp
           *    37727    41203: contig of 3477 bp in length
           *    41204    41303: gap of 100 bp
           *    41304    44513: contig of 3210 bp in length
           *    44514    44613: gap of 100 bp
           *    44614    48707: contig of 4094 bp in length
           *    48708    48807: gap of 100 bp
           *    48808    54126: contig of 5319 bp in length
           *    54127    54226: gap of 100 bp
           *    54227    62040: contig of 7814 bp in length
           *    62041    62140: gap of 100 bp
           *    62141    71155: contig of 9015 bp in length
           *    71156    71255: gap of 100 bp
           *    71256    80764: contig of 9509 bp in length
           *    80765    80864: gap of 100 bp
           *    80865    91109: contig of 10245 bp in length
           *    91110    91209: gap of 100 bp
           *    91210   101437: contig of 10228 bp in length
           *   101438   101537: gap of 100 bp
           *   101538   121402: contig of 19865 bp in length
           *   121403   121502: gap of 100 bp
           *   121503   144542: contig of 23040 bp in length.
FEATURES             Location/Qualifiers
     source          1. .144542
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-123N2"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1042
                     /note="assembly_fragment"
     gap             1043. .1142
                     /estimated_length=100
     misc_feature    1143. .2215
                     /note="assembly_fragment"
     gap             2216. .2315
                     /estimated_length=100
     misc_feature    2316. .3520
                     /note="assembly_fragment"
     gap             3521. .3620
                     /estimated_length=100
     misc_feature    3621. .6235
                     /note="assembly_fragment"
     gap             6236. .6335
                     /estimated_length=100
     misc_feature    6336. .8261
                     /note="assembly_fragment"
     gap             8262. .8361
                     /estimated_length=100
     misc_feature    8362. .10203
                     /note="assembly_fragment"
     gap             10204. .10303
                     /estimated_length=100
     misc_feature    10304. .11741
                     /note="assembly_fragment"
     gap             11742. .11841
                     /estimated_length=100
     misc_feature    11842. .13740
                     /note="assembly_fragment"
     gap             13741. .13840
                     /estimated_length=100
     misc_feature    13841. .16452
                     /note="assembly_fragment"
     gap             16453. .16552
                     /estimated_length=100
     misc_feature    16553. .19320
                     /note="assembly_fragment"
     gap             19321. .19420
                     /estimated_length=100

  Query Match             1.7%;  Score 52;  DB 14;  Length 144542;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110687 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 110738


RESULT 115
AC110281
LOCUS       AC110281               146199 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 11 clone CTD-2132H1 map 11, 6 ordered
            pieces.
ACCESSION   AC110281
VERSION     AC110281.4  GI:23499587
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 146199)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone CTD-2132H1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 146199)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
```

```
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 146199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 4, 2002 this sequence version replaced gi:23343843.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L23986
                Center clone name: 2132_H_1
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    30921: contig of 30921 bp in length
            *    30922    31021: gap of 100 bp
            *    31022    67623: contig of 36602 bp in length
            *    67624    67723: gap of 100 bp
            *    67724    88585: contig of 20862 bp in length
            *    88586    88685: gap of 100 bp
            *    88686    93649: contig of 4964 bp in length
            *    93650    93749: gap of 100 bp
            *    93750    98725: contig of 4976 bp in length
            *    98726    98825: gap of 100 bp
            *    98826   146199: contig of 47374 bp in length.
FEATURES             Location/Qualifiers
     source          1. .146199
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="CTD-2132H1"
                     /clone_lib="CITD1 Human BAC"
     gap             30922. .31021
                     /estimated_length=100
     gap             67624. .67723
                     /estimated_length=100
     gap             88586. .88685
                     /estimated_length=100
     gap             93650. .93749
                     /estimated_length=100
     gap             98726. .98825
                     /estimated_length=100
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 146199;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10626 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 10677


RESULT 116
AC009247
LOCUS       AC009247              146376 bp    DNA     linear   PRI 14-DEC-2001
DEFINITION  Homo sapiens 3q27  BAC RP11-410F19 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC009247
VERSION     AC009247.13  GI:17737024
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 146376)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Garcia,D.K., Gorrell,J.H.,
            Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
            Hodgson,A., Hogues,M., Holloway,C., Hosak,H., Jackson,L.E.,
            Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondejewski,N., Kong,Y.,
            Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,
            Lozado,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P.,
            Mei,G., Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S.,
            Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B.,
            Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D.,
            Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H.,
            Simon,M., Sparks,A., Stamps,A., Sucgang,R., Tabor,P., Taylor,T.,
            Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,
            Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 146376)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-1999) Human Genome Sequencing Center, Department
```

```
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 146376)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 146376)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 146376)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-DEC-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 14, 2001 this sequence version replaced gi:13489132.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

            QUALSTAT-REPORT.
FEATURES             Location/Qualifiers
     source          1. .146376
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-410F19"
     repeat_region   789. .825
                     /rpt_family="AT_rich"
     misc_feature    1371. .1711
                     /note="Region: Similar to Hs#S1074034 oz46g11.x1 Homo
                     sapiens cDNA, AI079997"
     repeat_region   1737. .1781
                     /rpt_family="(CA)n"
     repeat_region   3500. .3520
                     /rpt_family="AT_rich"
     misc_feature    join(4659. .4721,6288. .6513)
                     /note="Region: Similar to Hs#S1609703 DKFZp762L097_r1 Homo
                     sapiens cDNA, AL121466"
     repeat_region   5272. .5325
                     /rpt_family="GC_rich"
     repeat_region   6178. .6278
                     /rpt_family="GC_rich"
     repeat_region   6495. .6637
                     /rpt_family="C-rich"
     repeat_region   6695. .6718
                     /rpt_family="GC_rich"
     repeat_region   7005. .7142
                     /rpt_family="G-rich"
     repeat_region   8361. .8487
                     /rpt_family="MIR"
     repeat_region   9046. .9067
                     /rpt_family="(TTG)n"
     repeat_region   complement(9153. .9513)
                     /rpt_family="THE1B"
     repeat_region   complement(10270. .10517)
                     /rpt_family="AluJo"
     repeat_region   complement(10736. .11062)
                     /rpt_family="MER1B"
     repeat_region   11101. .11155
                     /rpt_family="MIR"
     repeat_region   11362. .11452
                     /rpt_family="L1MB8"
     repeat_region   complement(11696. .11997)
                     /rpt_family="L2"
     repeat_region   12202. .12827
                     /rpt_family="HAL1"
     repeat_region   12980. .13267
                     /rpt_family="AluJo"
     repeat_region   13473. .13671
                     /rpt_family="L2"
     repeat_region   complement(14040. .14312)
                     /rpt_family="AluJb"
     repeat_region   complement(14495. .14807)
                     /rpt_family="AluSp"
     repeat_region   complement(14897. .15005)
                     /rpt_family="MER5A"
     repeat_region   complement(15090. .15172)
                     /rpt_family="L1MD3"
     repeat_region   complement(15182. .15221)
                     /rpt_family="L1MC4"
     repeat_region   15222. .15533
                     /rpt_family="AluSq"
     repeat_region   complement(15534. .15650)
                     /rpt_family="L1MC4"
     repeat_region   15651. .15954
                     /rpt_family="AluYa5"
     repeat_region   complement(15955. .16485)
                     /rpt_family="L1MC4"
     repeat_region   complement(16586. .16678)
                     /rpt_family="AluSq/x"
     repeat_region   complement(16679. .16966)
                     /rpt_family="L1MC4"
     repeat_region   17043. .17351
                     /rpt_family="AluSg1"
     misc_feature    17320. .17410
                     /function="Low Coverage"
     repeat_region   17376. .17407
                     /rpt_family="(TTTTG)n"
     repeat_region   complement(17408. .17699)
                     /rpt_family="AluSx"
     repeat_region   complement(17700. .17843)
                     /rpt_family="L1MC4"
     repeat_region   complement(17878. .18184)
                     /rpt_family="AluSg"
     repeat_region   complement(18422. .18733)
                     /rpt_family="AluY"
     repeat_region   complement(18919. .20097)
```

```
                     /rpt_family="L1ME"
     repeat_region   20107. .20275
                     /rpt_family="L1PA13"
     repeat_region   20276. .20580
                     /rpt_family="AluSx"
     repeat_region   20581. .20741
                     /rpt_family="L1PA13"
     repeat_region   20758. .20865
                     /rpt_family="HAL1"
     repeat_region   20879. .20913
                     /rpt_family="(TTTTA)n"
     repeat_region   complement(20914. .21210)
                     /rpt_family="AluSg"
     repeat_region   complement(21253. .21311)
                     /rpt_family="LTR33"
     repeat_region   complement(21334. .21422)
                     /rpt_family="MER2"

  Query Match             1.7%;  Score 52;  DB 8;  Length 146376;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37189 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 37240


RESULT 117
CNS01DTX/c
LOCUS       CNS01DTX              146740 bp    DNA     linear   PRI 28-JUN-2001
DEFINITION  Human chromosome 14 DNA sequence BAC C-2002H8 of library CalTech-D
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL132989
VERSION     AL132989.5  GI:14586417
KEYWORDS    HTG; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 146740)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 146740)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Jul 2, 2001 this sequence version replaced gi:12580632.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : R-349A22
            Downstream BAC (overlapping the SP6 end) : R-517O13 (AC=AL139021)
            -------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 8.73x in Q20 bases; sum-of-contigs
            --------------
            ----------------------
            Overall quality chart :
            Range     :   bases
            0         :
            1  -  9   :
            10 - 19   :       1
            20 - 29   :       2
            30 - 39   :      41
            40 - 49   :    1267
            50 - 59   :    3807
            60 - 69   :    6119
            70 - 79   :   18019
            80 - 89   :   48401
            90 - 99   :   69083
            ----------------------
            Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1. .146740
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="C-2002H8"
                     /clone_lib="CalTech-D"
     STS             30201. .30314
                     /note="matching EMBL:G20444
                     RHdb:RH17151
                     RHdb:RH16912
                     dbSTS:STS11150
                     Identified using the e-PCR software (G. Schuler)"
     STS             30335. .30431
                     /note="matching EMBL:Z38484
                     RHdb:RH91710
                     dbSTS:STS64724
                     Identified using the e-PCR software (G. Schuler)"
     STS             61096. .61238
                     /note="matching EMBL:N22004
                     RHdb:RH69274
                     dbSTS:STS49152
                     Identified using the e-PCR software (G. Schuler)"
     STS             61230. .61323
                     /note="matching EMBL:G38638
                     RHdb:RH97264
                     RHdb:RH83920
                     dbSTS:STS59275
                     Identified using the e-PCR software (G. Schuler)"
     STS             63152. .63253
                     /note="matching EMBL:R85973
                     RHdb:RH53753
                     RHdb:RH14490
                     dbSTS:STS1946
                     Identified using the e-PCR software (G. Schuler)"
     STS             67540. .67664
                     /note="matching EMBL:T63600
                     RHdb:RH53926
                     RHdb:RH14889
                     dbSTS:STS8565
                     Identified using the e-PCR software (G. Schuler)"
     STS             67545. .67664
                     /note="matching EMBL:G28318
                     RHdb:RH33309
                     dbSTS:STS8564
                     Identified using the e-PCR software (G. Schuler)"
     STS             74562. .74829
                     /note="matching EMBL:H95046
                     RHdb:RH68861
                     dbSTS:STS48768
                     Identified using the e-PCR software (G. Schuler)"
     STS             96307. .96430
                     /note="matching EMBL:G30132
                     RHdb:RH37908
                     dbSTS:STS22657
                     Identified using the e-PCR software (G. Schuler)"
     STS             134360. .134499
                     /note="matching EMBL:G33047
                     RHdb:RH67742
                     dbSTS:STS47669
                     Identified using the e-PCR software (G. Schuler)"
ORIGIN
```

```
  Query Match             1.7%;  Score 52;  DB 8;  Length 146740;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136103 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 136052


RESULT 118
HSA300188/c
LOCUS       HSA300188             148049 bp    DNA     linear   PRI 15-APR-2005
DEFINITION  Homo sapiens MASP-2 gene (partial), PM-scl gene, FRAP2 gene and
            CDT6 gene, clone RPCI.11-99P18.
ACCESSION   AJ300188
VERSION     AJ300188.1  GI:11136974
KEYWORDS    CDT6 protein; FRAP2 gene; mannan binding lectin associated serine
            protease-2; MASP-2 protein; PM-scl autoantigen; PM-scl gene;
            rapamycin associated protein FRAP2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Stover,C., Endo,Y., Takahashi,M., Lynch,N.J., Constantinescu,C.,
            Vorup-Jensen,T., Thiel,S., Friedl,H., Hankeln,T., Hall,R.,
            Gregory,S., Fujita,T. and Schwaeble,W.
  TITLE     The human gene for mannan-binding lectin-associated serine
            protease-2 (MASP-2), the effector component of the lectin route of
            complement activation, is part of a tightly linked gene cluster on
            chromosome 1p36.2-3
  JOURNAL   Genes Immun. 2 (3), 119-127 (2001)
   PUBMED   11426320
REFERENCE   2  (bases 1 to 148049)
  AUTHORS   Friedl,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2000) Friedl H., Genomics Services, GENterprise
            GmbH, Becherweg 32, D-55099 Mainz, GERMANY
FEATURES             Location/Qualifiers
     source          1. .148049
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1p36.2-3"
                     /clone="BAC clone 99P18"
                     /clone_lib="Human RPCI.11"
     gene            complement(<3563. .>7906)
                     /gene="MASP-2"
     mRNA            complement(join(<3563. .4679,7203. .7277,7772. .7906))
                     /gene="MASP-2"
     CDS             complement(join(3916. .4679,7203. .7277,7772. .>7906))
                     /gene="MASP-2"
                     /function="cleavage of complement component C4"
                     /codon_start=3
                     /product="Masp-2 protein"
                     /protein_id="CAC15568.1"
                     /db_xref="GI:11136975"
                     /db_xref="GOA:Q96QG4"
                     /db_xref="InterPro:IPR000436"
                     /db_xref="InterPro:IPR001254"
                     /db_xref="UniProt/TrEMBL:Q96QG4"
                     /translation="VDCGPPDDLPSGRAEYITGPGVTTYKAVIQYSCEETFYTMKVND
                     GKYVCEADGFWTSSKGEKSLPVCEPVCGLSARTTGGRIYGGQKAKPGDFPWQVLILGG
                     TTAAGALLYDNWVLTAAHAVYEQKHDASALDIRMGTLKRLSPHYTQAWSEAVFIHEGY
                     THDAGFDNDIALIKLNNKVVINSNITPICLPRKEAESFMRTDDIGTASGWGLTQRGFL
                     ARNLMYVDIPIVDHQKCTAAYEKPPYPRGSVTANMLCAGLESGGKDSCRGDSGGALVF
                     LDSETERWFVGGIVSWGSMNCGEAGQYGVYTKVINYIPWIENIISDF"
     intron          complement(4680. .7202)
                     /gene="MASP-2"
                     /number=2
     exon            complement(7203. .7277)
                     /gene="MASP-2"
                     /number=2
     intron          complement(7278. .7771)
                     /gene="MASP-2"
                     /number=1
     exon            complement(7772. .>7906)
                     /gene="MASP-2"
                     /number=1
     gene            complement(15104. .48352)
                     /gene="PM-scl"
     mRNA            complement(join(15104. .15231,16491. .16567,17128. .17189,
                     18042. .18213,19383. .19456,20570. .20654,22715. .22810,
                     25329. .25435,25852. .25930,26088. .26138,28198. .28309,
                     28986. .29036,29249. .29397,29567. .29723,31173. .31363,
                     35933. .36076,36285. .36395,36608. .36683,39037. .39151,
                     39497. .39662,39976. .40080,44241. .44364,46503. .46639,
                     48204. .48352))
                     /gene="PM-scl"
     exon            complement(15104. .15231)
                     /gene="PM-scl"
                     /number=24
     CDS             complement(join(15201. .15231,16491. .16567,17128. .17189,
                     18042. .18213,19383. .19456,20570. .20654,22715. .22810,
                     25329. .25435,25852. .25930,26088. .26138,28198. .28309,
                     28986. .29036,29249. .29397,29567. .29723,31173. .31363,
                     35933. .36076,36285. .36395,36608. .36683,39037. .39151,
                     39497. .39662,39976. .40080,44241. .44364,46503. .46639,
                     48204. .48314))
                     /gene="PM-scl"
                     /codon_start=1
                     /product="PM-scl autoantigen"
                     /protein_id="CAC15569.1"
                     /db_xref="GI:11136976"
                     /db_xref="GOA:Q01780"
                     /db_xref="UniProt/Swiss-Prot:Q01780"
                     /translation="MAPPSTREPRVLSATSATKSDGEMVLPGFPDADSFVKFALGSVV
                     AVTKASGGLPQFGDEYDFYRSFPGFQAFCETQGDRLLQCMSRVMQYHGCRSNIKDRSK
                     VTELEDKFDLLVDANDVILERVGILLDEASGVNKNQQPVLPAGLQVPKTVVSSWNRKA
                     AEYGKKAKSETFRLLHAKNIIRPQLKFREKIDNSNTPFLPKIFIKPNAQKPLPQALSK
                     ERRERPQDRPEDLDVPPALADFIHQQRTQQVEQDMFAHPYQYELNHFTPADAVLQKPQ
                     PQLYRPIEETPCHFISSLDELVELNEKLLNCQEFAVDLEHHSYRSFLGLTCLMQISTR
                     TEDFIIDTLELRSDMYILNESLTDPAIVKVFHGADSDIEWLQKDFGLYVVNMFDTHQA
                     ARLLNLGRHSLDHLLKLYCNVDSNKQYQLADWRIRPLPEEMLSYARDDTHYLLYIYDK
                     MRLEMWERGNGQPVQLQVVWQRSRDICLKKFIKPIFTDESYLELYRKQKKHLNTQQLT
                     AFQLLFAWRDKTARREDESYGYVLPNHMMLKIAEELPKEPQGIIACCNPVPPLVRQQI
                     NEMHLLIQQAREMPLLKSEVAAGVKKSGPLPSAERLENVLFGPHDCSHAPPDGYPIIP
                     TSGSVPVQKQASLFPDEKEDNLLGTTCLIATAVITLFNEPSAEDSKKGPLTVAQKKAQ
                     NIMESFENPFRMISNRWKLAQVQVQKDSKEAVKKKAAEQTAAREQAKEACKAAAEQAI
                     SVRQQVVLENAAKKRERATSDPRTTEQKQEKKRLKISKKPKDPEPPEKEFTPYDYSQS
                     DFKAFAGNSKSKVSSQFDPNKQTPSGKKCIAAKKIKQSVGNKSMSFPTGKSDRGFRYN
                     WPQR"
     intron          complement(15232. .16490)
                     /gene="PM-scl"
                     /number=23
     exon            complement(16491. .16567)
                     /gene="PM-scl"
                     /number=23
     intron          complement(16568. .17127)
                     /gene="PM-scl"
                     /number=22
     exon            complement(17128. .17189)
                     /gene="PM-scl"
                     /number=22
     intron          complement(17190. .18041)
                     /gene="PM-scl"
                     /number=21
     exon            complement(18042. .18213)
                     /gene="PM-scl"
                     /number=21
     intron          complement(18214. .19382)
                     /gene="PM-scl"
                     /number=20
```

```
     exon            complement(19383. .19456)
                     /gene="PM-scl"
                     /number=20
     intron          complement(19457. .20569)
                     /gene="PM-scl"
                     /number=19
     exon            complement(20570. .20654)
                     /gene="PM-scl"
                     /number=19
     intron          complement(20655. .22714)
                     /gene="PM-scl"
                     /number=18
     exon            complement(22715. .22810)
                     /gene="PM-scl"
                     /number=18
     intron          complement(22811. .25328)
                     /gene="PM-scl"
                     /number=17
     exon            complement(25329. .25435)
                     /gene="PM-scl"
                     /number=17
     intron          complement(25436. .25851)
                     /gene="PM-scl"
                     /number=16
     exon            complement(25852. .25930)
                     /gene="PM-scl"
                     /number=16
     intron          complement(25931. .26087)
                     /gene="PM-scl"
                     /number=15
     exon            complement(26088. .26138)
                     /gene="PM-scl"
                     /number=15
     intron          complement(26139. .28197)
                     /gene="PM-scl"
                     /number=14
     exon            complement(28198. .28309)
                     /gene="PM-scl"
                     /number=14
     intron          complement(28310. .28985)
                     /gene="PM-scl"
                     /number=13
     exon            complement(28986. .29036)
                     /gene="PM-scl"
                     /number=13
     intron          complement(29037. .29248)
                     /gene="PM-scl"
                     /number=12
     exon            complement(29249. .29397)
                     /gene="PM-scl"
                     /number=12
     intron          complement(29398. .29566)
                     /gene="PM-scl"
                     /number=11
     exon            complement(29567. .29723)
                     /gene="PM-scl"
                     /number=11
     intron          complement(29724. .31172)
                     /gene="PM-scl"
                     /number=10
     exon            complement(31173. .31363)
                     /gene="PM-scl"
                     /number=10
     intron          complement(31364. .35932)
                     /gene="PM-scl"
                     /number=9

  Query Match             1.7%;  Score 52;  DB 8;  Length 148049;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70758 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 70707

RESULT 119
AC002351/c
LOCUS       AC002351               149679 bp    DNA     linear   PRI 21-FEB-2003
DEFINITION  Homo sapiens 12 PAC RP1-46F2 (Roswell Park Cancer Institute Human
            PAC Library) complete sequence.
ACCESSION   AC002351
VERSION     AC002351.2  GI:28460712
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 149679)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 149679)
  AUTHORS   Chiu,M.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1997) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   7  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   8  (bases 1 to 149679)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Feb 21, 2003 this sequence version replaced gi:2811094.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
            ml.
FEATURES             Location/Qualifiers
     source          1. .149679
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP1-46F2"
     misc_feature    1. .220
                     /note="overlaps bases 102372. .102591 of clone AC002978"
                     /function="clone overlap"
     repeat_region   36. .62
                     /rpt_family="(TCTG)n"
     repeat_region   763. .879
                     /rpt_family="CT-rich"
     repeat_region   1019. .1056
                     /rpt_family="(CAT)n"
     repeat_region   complement(1124. .1203)
                     /rpt_family="MIR"
     repeat_region   complement(1558. .1748)
                     /rpt_family="MIR"
     repeat_region   2367. .2533
                     /rpt_family="(CA)n"
     repeat_region   3148. .3270
                     /rpt_family="MIR"
     repeat_region   3420. .3478
                     /rpt_family="MIR"
     repeat_region   complement(3479. .3560)
                     /rpt_family="AluJ/FLAM"
     repeat_region   3562. .3653
                     /rpt_family="MIR"
     repeat_region   complement(3772. .3831)
                     /rpt_family="L2"
     repeat_region   complement(3795. .3881)
                     /rpt_family="MIR"
     repeat_region   complement(3882. .4167)
                     /rpt_family="AluJo"
     repeat_region   complement(4168. .4291)
                     /rpt_family="MIR"
     repeat_region   4853. .4933
                     /rpt_family="MIR"
     repeat_region   4934. .5236
                     /rpt_family="AluSx"
     repeat_region   5237. .5256
                     /rpt_family="MIR"
     repeat_region   5257. .5532
                     /rpt_family="MLT1H"
     repeat_region   complement(6419. .6716)
                     /rpt_family="AluJo"
     repeat_region   6752. .7164
                     /rpt_family="L1ME3"
     repeat_region   7165. .7469
                     /rpt_family="AluJb"
     repeat_region   7470. .7684
                     /rpt_family="L1ME3"
     repeat_region   7739. .8023
                     /rpt_family="AluSq"
     repeat_region   8024. .8054
                     /rpt_family="(CAAA)n"
     repeat_region   8055. .8078
                     /rpt_family="AluSq"
     repeat_region   8081. .8370
                     /rpt_family="AluSc"
     repeat_region   8371. .8402
                     /rpt_family="(CAAA)n"
     repeat_region   8663. .8725
                     /rpt_family="MIR"
     repeat_region   8845. .8995
                     /rpt_family="CT-rich"
     repeat_region   complement(9002. .9274)
                     /rpt_family="AluJo"

  Query Match             1.7%;  Score 52;  DB 8;  Length 149679;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45739 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 45688


RESULT 120
AC012070
LOCUS       AC012070              149776 bp    DNA     linear   PRI 21-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-318K6 from 2, complete sequence.
```

```
ACCESSION   AC012070
VERSION     AC012070.8  GI:21747558
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 149776)
  AUTHORS   Isak,A. and Abbott,A.
  TITLE     The sequence of Homo sapiens BAC clone RP11-318K6
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 149776)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 149776)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 149776)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 149776)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 149776)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 149776)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 14, 2002 this sequence version replaced gi:13162544.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0318K06
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-335M11, 200 bp overlap; the
            clone sequenced to the right is RP11-674L1.  Actual start of this
            clone is at base position 150598 of RP11-335M11; actual end is at
            base position 149776 of RP11-318K6.

            There are polymorphic base pair differences in the overlap between
            the clone RP11-318K6 and RP11-674L1.
FEATURES             Location/Qualifiers
     source          1. .149776
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-318K6"
                     /clone_lib="RPCI-11"
     misc_feature    9331. .9578
                     /note="CpG_island (%GC=60.1, o/e=1.05, #CpGs=23)"
     misc_feature    28997. .29214
                     /note="CpG_island (%GC=61.0, o/e=0.86, #CpGs=19)"
     misc_feature    52878. .53098
                     /note="CpG_island (%GC=60.2, o/e=0.90, #CpGs=18)"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 149776;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81593 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 81644


RESULT 121
AC018791
LOCUS       AC018791              150159 bp    DNA     linear   HTG 01-APR-2000
DEFINITION  Homo sapiens clone RP11-31D10, WORKING DRAFT SEQUENCE, 26 unordered
            pieces.
ACCESSION   AC018791
VERSION     AC018791.3  GI:7382241
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 150159)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-31D10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150159)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
            O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
            Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
```

```
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
            and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 1, 2000 this sequence version replaced gi:6692294.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L4923
                Center clone name: 31_D_10
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 137932 bases at least Q40
                Consensus quality: 143842 bases at least Q30
                Consensus quality: 146083 bases at least Q20
                Insert size: 169000; agarose-fp
                Insert size: 147659; sum-of-contigs
                Quality coverage: 4.2 in Q20 bases; agarose-fp
                Quality coverage: 4.8 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 26 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1239: contig of 1239 bp in length
            *     1240     1339: gap of 100 bp
            *     1340     3184: contig of 1845 bp in length
            *     3185     3284: gap of 100 bp
            *     3285     5534: contig of 2250 bp in length
            *     5535     5634: gap of 100 bp
            *     5635     8315: contig of 2681 bp in length
            *     8316     8415: gap of 100 bp
            *     8416    10195: contig of 1780 bp in length
            *    10196    10295: gap of 100 bp
            *    10296    13456: contig of 3161 bp in length
            *    13457    13556: gap of 100 bp
            *    13557    15457: contig of 1901 bp in length
            *    15458    15557: gap of 100 bp
            *    15558    17848: contig of 2291 bp in length
            *    17849    17948: gap of 100 bp
            *    17949    19893: contig of 1945 bp in length
            *    19894    19993: gap of 100 bp
            *    19994    23457: contig of 3464 bp in length
            *    23458    23557: gap of 100 bp
            *    23558    26795: contig of 3238 bp in length
            *    26796    26895: gap of 100 bp
            *    26896    30999: contig of 4104 bp in length
            *    31000    31099: gap of 100 bp
            *    31100    35140: contig of 4041 bp in length
            *    35141    35240: gap of 100 bp
            *    35241    39261: contig of 4021 bp in length
            *    39262    39361: gap of 100 bp
            *    39362    42633: contig of 3272 bp in length
            *    42634    42733: gap of 100 bp
            *    42734    47924: contig of 5191 bp in length
            *    47925    48024: gap of 100 bp
            *    48025    52442: contig of 4418 bp in length
            *    52443    52542: gap of 100 bp
            *    52543    59093: contig of 6551 bp in length
            *    59094    59193: gap of 100 bp
            *    59194    67470: contig of 8277 bp in length
            *    67471    67570: gap of 100 bp
            *    67571    74426: contig of 6856 bp in length
            *    74427    74526: gap of 100 bp
            *    74527    87418: contig of 12892 bp in length
            *    87419    87518: gap of 100 bp
            *    87519    95952: contig of 8434 bp in length
            *    95953    96052: gap of 100 bp
            *    96053   106060: contig of 10008 bp in length
            *   106061   106160: gap of 100 bp
            *   106161   121730: contig of 15570 bp in length
            *   121731   121830: gap of 100 bp
            *   121831   135347: contig of 13517 bp in length
            *   135348   135447: gap of 100 bp
            *   135448   150159: contig of 14712 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150159
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-31D10"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1239
                     /note="assembly_fragment"
     gap             1240. .1339
                     /estimated_length=100
     misc_feature    1340. .3184
                     /note="assembly_fragment"
     gap             3185. .3284
                     /estimated_length=100
     misc_feature    3285. .5534
                     /note="assembly_fragment"
     gap             5535. .5634
                     /estimated_length=100
     misc_feature    5635. .8315
                     /note="assembly_fragment"
     gap             8316. .8415
                     /estimated_length=100
     misc_feature    8416. .10195
                     /note="assembly_fragment"
     gap             10196. .10295
                     /estimated_length=100
     misc_feature    10296. .13456
                     /note="assembly_fragment"
     gap             13457. .13556
                     /estimated_length=100
     misc_feature    13557. .15457
                     /note="assembly_fragment"
     gap             15458. .15557
                     /estimated_length=100
     misc_feature    15558. .17848
                     /note="assembly_fragment"
     gap             17849. .17948
                     /estimated_length=100
     misc_feature    17949. .19893
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             19894. .19993
                     /estimated_length=100
     misc_feature    19994. .23457
                     /note="assembly_fragment"
     gap             23458. .23557
                     /estimated_length=100
     misc_feature    23558. .26795
                     /note="assembly_fragment"
     gap             26796. .26895
                     /estimated_length=100
     misc_feature    26896. .30999
                     /note="assembly_fragment"
     gap             31000. .31099
                     /estimated_length=100
```

```
     misc_feature    31100. .35140
                     /note="assembly_fragment"
     gap             35141. .35240
                     /estimated_length=100
     misc_feature    35241. .39261
                     /note="assembly_fragment"
     gap             39262. .39361
                     /estimated_length=100
     misc_feature    39362. .42633
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             42634. .42733
                     /estimated_length=100
     misc_feature    42734. .47924
                     /note="assembly_fragment"
     gap             47925. .48024
                     /estimated_length=100
     misc_feature    48025. .52442
                     /note="assembly_fragment"
     gap             52443. .52542
                     /estimated_length=100
     misc_feature    52543. .59093
                     /note="assembly_fragment"

  Query Match             1.7%;  Score 52;  DB 14;  Length 150159;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122542 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 122593


RESULT 122
CS086345
LOCUS       CS086345              150437 bp    DNA     linear   PAT 25-MAY-2005
DEFINITION  Sequence 50 from Patent WO2005042786.
ACCESSION   CS086345
VERSION     CS086345.1  GI:66711917
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Harris,C. and Davis,L.
  TITLE     Compositions and methods for glioma classification
  JOURNAL   Patent: WO 2005042786-A 50 12-MAY-2005;
            Exagen Diagnostics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .150437
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 150437;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125751 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 125802


RESULT 123
AL138724
LOCUS       AL138724              150846 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-500C11 on chromosome 6 Contains
            the 5' end of the NUP153 gene for nucleoporin 153kD, the 3' end of
            a gene for a kinesin-like protein (RBKIN) and CpG islands, complete
            sequence.
ACCESSION   AL138724
VERSION     AL138724.12  GI:13234804
KEYWORDS    HTG; CpG island; kinesin; nucleoporin; NUP153; RBKIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 150846)
  AUTHORS   Griffiths,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 5, 2001 this sequence version replaced gi:13161550.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-500C11 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .150846
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-500C11"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-500C11"
     gene            join(complement(8974. .9084),
                     complement(AL157776.15:1295. .61258),
                     complement(AL138824.19:<83780. .94906))
                     /gene="NUP153"
                     /locus_tag="RP11-14H3.6-001"
     mRNA            join(complement(8974. .9084),
                     complement(AL157776.15:61036. .61258),
                     complement(AL157776.15:48162. .48410),
                     complement(AL157776.15:47869. .48008),
                     complement(AL157776.15:47545. .47673),
                     complement(AL157776.15:42070. .42186),
                     complement(AL157776.15:41933. .41977),
                     complement(AL157776.15:41615. .41668),
                     complement(AL157776.15:37879. .38025),
                     complement(AL157776.15:34658. .34710),
                     complement(AL157776.15:34293. .34419),
                     complement(AL157776.15:21803. .21940),
                     complement(AL157776.15:20447. .20545),
                     complement(AL157776.15:18707. .18794),
                     complement(AL157776.15:12579. .12704),
                     complement(AL157776.15:9793. .10410),
```

```
                     complement(AL157776.15:5290. .5484),
                     complement(AL157776.15:1295. .2179),
                     complement(AL138824.19:94550. .94906),
                     complement(AL138824.19:93303. .93575),
                     complement(AL138824.19:85269. .85437),
                     complement(AL138824.19:<83780. .84923))
                     /gene="NUP153"
                     /locus_tag="RP11-14H3.6-001"
                     /product="nucleoporin 153kDa"
                     /note="match: ESTs: AA020948 AA280626 AA281956 AA287143
                     AA647494 AA987890 AI373182 AI392741 AI568716 AI675749
                     AI809345 AI863498 AI918052 AL040791 AL043770 AL044844
                     AL597150 AU125074 AU128712 AU137373 AU137888 AU148810
                     AW172435 AW269393 AW269401 AW269570 AW502128 AW502296
                     AW604841 AW859945 AW859990 AW937856 AW997580 AW997597
                     BB529859 BE085871 BE333864 BE381487 BE539035 BE539837
                     BE885083 BE886640 BE909229 BE930529 BF120400 BF135241
                     BF212928 BF305010 BF702538 BF732659 BF782907 BF843418
                     BF844739 BF934823 BF934824 BF991876 BF991889 BF992423
                     BG067610 BG080749 BG106248 BG256869 BG259567 BG387515
                     BG505279 BG986900 BI040158 BI040160 BI040168 BI052625
                     BM090763 F13065 H89627 L44346 R17048 T09409 T75378
                     match: cDNAs: AF045567.1 AF434196.1 L06821.1 Z25535.1"
CDS                  join(complement(8974. .9084),
                     complement(AL157776.15:61036. .61258),
                     complement(AL157776.15:48162. .48410),
                     complement(AL157776.15:47869. .48008),
                     complement(AL157776.15:47545. .47673),
                     complement(AL157776.15:42070. .42186),
                     complement(AL157776.15:41933. .41977),
                     complement(AL157776.15:41615. .41668),
                     complement(AL157776.15:37879. .38025),
                     complement(AL157776.15:34658. .34710),
                     complement(AL157776.15:34293. .34419),
                     complement(AL157776.15:21803. .21940),
                     complement(AL157776.15:20447. .20545),
                     complement(AL157776.15:18707. .18794),
                     complement(AL157776.15:12579. .12704),
                     complement(AL157776.15:9793. .10410),
                     complement(AL157776.15:5290. .5484),
                     complement(AL157776.15:1295. .2179),
                     complement(AL138824.19:94550. .94906),
                     complement(AL138824.19:93303. .93575),
                     complement(AL138824.19:85269. .85437),
                     complement(AL138824.19:<84842. .84923))
                     /gene="NUP153"
                     /locus_tag="RP11-14H3.6-001"
                     /standard_name="OTTHUMP00000016070"
                     /note="match: proteins: O73749 P49790 P49791"
                     /codon_start=1
                     /product="nucleoporin 153kDa"
                     /protein_id="CAI40945.1"
                     /db_xref="GI:57208789"
                     /db_xref="GOA:Q5TAZ2"
                     /db_xref="InterPro:IPR000194"
                     /db_xref="InterPro:IPR001876"
                     /db_xref="UniProt/TrEMBL:Q5TAZ2"
                     /translation="MASGAGGVGGGGGGKIRTRRCHQGPIKPYQQGRQQHQGILSRVT
                     ESVKNIVPGWLQRYFNKNEDVCSCSTDTSEVPRWPENKEDHLVYADEESSNITDGRIT
                     PEPAVSNTEEPSTTSTASNYPDVLTRPSLHRSHLNFSMLESPALHCQPSTSSAFPIGS
                     SGFSLVKEIKDSTSQHDDDNISTTSGFSSRASDKDITVSKNTSLPPLWSPEAERSHSL
                     SQHTATSSKKPAFNLSAFGTLSPSLGNSSILKTSQLGDSPFYPGKTTYGGAAAAVRQS
                     KLRNTPYQAPVRRQMKAKQLSAQSYGVTSSTARRILQSLEKMSSPLADAKRIPSIVSS
                     PLNSPLDRSGIDITDFQAKREKVDSQYPPVQRLMTPKPVSIATNRSVYFKPSLTPSGE
                     FRKTNQRIDNKCSTGYEKNMTPGQNREQRESGFSYPNFSLPAANGLSSGVGGGGGKMR
                     RERTRFVASKPLEEEEMEVPVLPKISLPITSSSLPTFNFSSPEITTSSPSPINSSQAL
                     TNKVQMTSPSSTGSPMFKFSSPIVKSTEANVLPPSSIGFTFSVPVAKTAELSGSSSTL
                     EPIISSSAHHVTTVNSTNCKKTPPEDCEGPFRPAEILKEGSVLDILKSPGFASPKIDS
                     VAAQPTATSPVVYTRPAISSFSSSGIGFGESLKAGSSWQCDTCLLQNKVTDNKCIACQ
                     AAKLSPRDTAKQTGIETPNKSGKTTLSASGTGFGDKFKPVIGTWDCDTCLVQNKPEAI
                     KCVACETPKPGTCVKRALTLTVVSESAETMTASSSSCTVTTGTLGFGDKFKRPIGSWE
                     CSVCCVSNNAEDNKCVSCMSEKPGSSVPASSSSTVPVSLPSGGSLGLEKFKKPEGSWD
                     CELCLVQNKADSTKCLACESAKPGTKSGFKGFDTSSSSSNSAASSSFKFGVSSSSSGP
                     SQTLTSTGNFKFGDQGGFKIGVSSDSGSINPMSEGFKFSKPIGDFKFGVSSESKPEEV
                     KKDSKNDNFKFGLSSGLSNPVSLTPFQFGVSNLGQEEKKEELPKSSSAGFSFGTGVIN
                     STPAPANTIVTSENKSSFNLGTIETKSASVAPFTCKTSEAKKEEMPATKGGFSFGNVE
                     PASLPSASVFVLGRTEEKQQEPVTSTSLVFGKKADNEEPKCQPVFSFGNSEQTKDENS
                     SKSTFSFSMTKPSEKESEQPAKATFAFGAQTSTTADQGAAKPVFSFLNNSSSSSSTPA
                     TSAGGGIFGSSTSSSNPPVATFVFGQSSNPVSSSAFGNTAESSTSQSLLFSQDSKLAT
                     TSSTGTAVTPFVFGPGASSNNTTTSGFGFGATTTSSSAGSSFVFGTGPSAPSASPAFG
                     ANQTPTFGQSQGASQPNPPGFGSISSSTALFPTGSQPAPPTFGTVSSSSQPPVFGQQP
                     SQSAFGSGTTPNSSSAFQFGSSTTNFNFTNNSPSGVFTFGANSSTPAASAQPSGSGGF
                     PFNQSPAAFTVGSNGKNVFSSSGTSFSGRKIKTAVRRRK"
gene                 join(AL023807.6:28627. .28752,AL023807.6:29017. .29107,
                     AL023807.6:117981. .117993,AL023807.6:142724. .142784,
                     AL023807.6:160039. .160131,AL023807.6:160544. .160724,
                     AL023807.6:164119. .164206,AL023807.6:165704. .165838,
                     AL023807.6:166672. .166784,complement(140169. .140280),
                     complement(139575. .139787),complement(136658. .136768),
                     complement(133798. .133932),complement(130937. .131067),
                     complement(128735. .128821),complement(128465. .128631),
                     complement(119717. .119930),complement(111465. .111627),
                     complement(108172. .108312),complement(107058. .107207),
                     complement(102649. .102810),complement(101963. .102136),
                     complement(99366. .99517),complement(97269. .97401),
                     complement(96946. .97092),complement(92569. .92607),
                     complement(90473. .90572),complement(88212. .88338),
                     complement(86343. .86398),complement(83874. .83998),
                     complement(83427. .83603),complement(82289. .82381),
                     complement(81640. .81796),complement(79974. .80041),
                     complement(77712. .77759),complement(76175. .76280),
                     complement(74605. .74756),complement(66392. .67643))
                     /locus_tag="RP11-500C11.2-002"
gene                 join(AL023807.6:28627. .28752,AL023807.6:29017. .29107,
                     AL023807.6:117981. .117993,AL023807.6:142724. .142784,
                     AL023807.6:160039. .160131,AL023807.6:160544. .160724,
                     AL023807.6:164119. .164206,AL023807.6:165704. .165838,
                     AL023807.6:166672. .166784,complement(140169. .140280),
                     complement(139575. .139787),complement(136658. .136768),
                     complement(133798. .133932),complement(130937. .131067),
                     complement(128735. .128821),complement(128465. .128631),
                     complement(119717. .119930),complement(111465. .111627),
                     complement(108172. .108312),complement(107058. .107207),
                     complement(102649. .102810),complement(101963. .102136),
                     complement(99366. .99517),complement(97269. .97401),
                     complement(96946. .97092),complement(92569. .92607),
                     complement(90473. .90572),complement(88212. .88338),
                     complement(86343. .86398),complement(83874. .83998),
                     complement(83427. .83603),complement(82289. .82381),
                     complement(81640. .81796),complement(79974. .80041),
                     complement(77712. .77759),complement(76175. .76280),
                     complement(74605. .74756),complement(73811. .73915),
                     complement(66392. .67643))

  Query Match             1.7%;  Score 52;  DB 8;  Length 150846;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4478 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4529


RESULT 124
AL662800
LOCUS       AL662800              151228 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-249D20 on chromosome 6 contains
            the GNL1 gene for guanine nucleotide binding protein-like 1, the
            gene for CAT56 protein, the ABCF1 gene for ATP-binding cassette,
            subfamily F, member 1, the PPP1R10 gene for protein phosphatase 1,
            regulatory subunit 10, the MRPS18B gene for  mitochondrial
            ribosomal protein S18B, two novel proteins, a prothymosin alpha
            (PTMA) pseudogene and four CpG islands, complete sequence.
ACCESSION   AL662800
VERSION     AL662800.4  GI:18181758
KEYWORDS    HTG; ABC27; ABC50; ABCF1; ATP-binding cassette, subfamily F, member
```

```
            1; CAT56; CpG island; DKFZp564H0223; FB19; GNL1; guanine nucleotide
            binding protein-like 1; HSPC183; HSR1; Human S18a; mitochondrial
            ribosomal protein S18B; MRP-S18-2; MRPS18-2; MRPS18B; PNUTS;
            PPP1R10; protein phosphatase 1, regulatory subunit 10; prothymosin
            alpha; PTD017; PTMA; S18amt; TMSA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 151228)
  AUTHORS   Griffiths,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 16, 2002 this sequence version replaced gi:18070933.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-249D20 is
            from a CHORI-501 human bac - PGF cell line library VECTOR:
            pTARBAC2.1
            This sequence was generated from part of bacterial clone contigs
            constructed by the MHC Haplotype Consortium and collaborators.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .151228
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-249D20"
                     /clone_lib="CHORI-501"
     misc_feature    1
                     /note="Clone_left_end: XXbac-249D20"
     misc_feature    11118
                     /note="Clone_left_end: XXbac-27H4"
     misc_feature    11123
                     /note="Clone_right_end: XXbac-34K20"
     gene            complement(46052. .56431)
                     /gene="GNL1"
                     /locus_tag="XXbac-BPG249D20.1-001"
     mRNA            complement(join(46052. .46488,46871. .47011,47287. .47449,
                     47527. .47705,52642. .52836,53259. .53354,53525. .53732,
                     54750. .54821,54939. .55090,55230. .55366,55730. .55895,
                     56306. .56431))
                     /gene="GNL1"
                     /locus_tag="XXbac-BPG249D20.1-001"
                     /product="guanine nucleotide binding protein-like 1"
                     /note="match: cDNAs: AK026391.1 BC010298.1 BC013243.1
                     BC013959.1 L25665.1 X65026.1"
     CDS             complement(join(46247. .46488,46871. .47011,47287. .47449,
                     47527. .47705,52642. .52836,53259. .53354,53525. .53732,
                     54750. .54821,54939. .55090,55230. .55366,55730. .55895,
                     56306. .56378))
                     /gene="GNL1"
                     /locus_tag="XXbac-BPG249D20.1-001"
                     /standard_name="OTTHUMP00000029186"
                     /note="match: proteins: P36916 Q91Z20 Q96CT5"
                     /codon_start=1
                     /product="guanine nucleotide binding protein-like 1"
                     /protein_id="CAI18152.1"
                     /db_xref="GI:55961299"
                     /db_xref="GOA:Q96CT5"
                     /db_xref="InterPro:IPR002917"
                     /db_xref="InterPro:IPR005289"
                     /db_xref="UniProt/TrEMBL:Q96CT5"
                     /translation="MPRKKPFSVKQKKKQLQDKRERKRGLQDGLRSSSNSRSGSRERR
                     EEQTDTSDGESVTHHIRRLNQQPSQGLGPRGYDPNRYRLHFERDSREEVERRKRAARE
                     QVLQPVSAELLELDIREVYQPGSVLDFPRRPPWSYEMSKEQLMSQEERSFQDYLGKIH
                     GAYSSEKLSYFEHNLETWRQLWRVLEMSDIVLLITDIRHPVVNFPPALYEYVTGELGL
                     ALVLVLNKVDLAPPALVVAWKHYFHQHYPQLHVVLFTSFPRDPRTPQDPSSVLKKSRR
                     RGRGWTRALGPEQLLRACEAITVGKVDLSSWREKIARDVAGATWGNGSGEEEEEEDGP
                     AVLVEQQTDSAMEPTGPTQERYKDGVVTIGCVGFPNVGKSSLINGLVGRKVVSVSRTP
                     GHTRYFQTYFLTPSVKLCDCPGLIFPSLLPRQLQVLAGIYPIAQIQEPYTAVGYLASR
                     IPVQALLHLRHPEAEDPSAEHPWCAWDICEAWAEKRGYKTAKAARNDVYRAANSLLRL
                     AVDGRLSLCFHPPGYSEQKGTWESHPETTELVVLQGRVGPAGDEEEEEEEELSSSCEE
                     EGEEDRDADEEGEGDEETPTSAPGSSLAGRNPYALLGEDEC"
     gene            join(57061. .57625,58325. .58387,61503. .61683,62009. .62299,
                     62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-003"
     gene            join(57061. .57625,58325. .58387,62009. .62299,62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-001"
     gene            join(57061. .57625,61503. .61683,62009. .62299,62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-002"
     mRNA            join(57061. .57625,58325. .58387,61503. .61683,62009. .62299,
                     62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-003"
                     /product="proline-rich polypeptide 3"
                     /note="match: ESTs: AL527443.1"
     mRNA            join(57061. .57625,58325. .58387,62009. .62299,62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-001"
                     /product="proline-rich polypeptide 3"
                     /note="match: ESTs: BI752625.1 BM545304.1 BM803167.1
                     match: cDNAs: U63336.1"
     mRNA            join(57061. .57625,61503. .61683,62009. .62299,62564. .63898)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-002"
                     /product="proline-rich polypeptide 3"
                     /note="match: ESTs: BM921809.1"
     CDS             join(57520. .57625,58325. .58387,62009. .62299,62564. .62670)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-001"
                     /standard_name="OTTHUMP00000029041"
                     /note="match: proteins: P79522 Q91YI5 Q96QB9"
                     /codon_start=1
                     /product="proline-rich polypeptide 3"
                     /protein_id="CAI18155.1"
                     /db_xref="GI:55961302"
                     /db_xref="GOA:Q96QB9"
                     /db_xref="InterPro:IPR000571"
                     /db_xref="UniProt/TrEMBL:Q96QB9"
                     /translation="MPKRKKQNHHQPPTQQQPPLPEREETGDEEDGSPIGPPSLLGPP
                     PMANGKPGDPKSALHRGPPGSRGPLIPPLLSLPPPPWGRGPIRRGLGPRSSPYGRGWW
                     GVNAEPPFPGPGHGGPTRGSFHKEQRNPRRLKSWSLIKNTCPPKDDPQVMEDKSDRPV
                     CRHFAKKGHCRYEDLCAFYHPGVNGPPL"
     CDS             join(57520. .57625,58325. .58387,61503. .61603)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-003"
                     /standard_name="OTTHUMP00000029043"
                     /note="match: proteins: P79522 Q91YI5 Q96QB9"
                     /codon_start=1
                     /product="proline-rich polypeptide 3"
                     /protein_id="CAI18154.1"
                     /db_xref="GI:55961301"
```

```
                     /db_xref="UniProt/TrEMBL:Q5STN7"
                     /translation="MPKRKKQNHHQPPTQQQPPLPEREETGDEEDGSPIGPPSLLGPP
                     PMANGKPGDPKSAPNWIPPTCPSSVEWINKLLCMHAMGLHCNENE"
     CDS             join(57520. .57625,61503. .61603)
                     /gene="PRR3"
                     /locus_tag="XXbac-BPG249D20.5-002"
                     /standard_name="OTTHUMP00000029042"
                     /note="match: proteins: P79522 Q91YI5 Q96QB9"
                     /codon_start=1
                     /product="proline-rich polypeptide 3"
                     /protein_id="CAI18153.1"
                     /db_xref="GI:55961300"
                     /db_xref="UniProt/TrEMBL:Q5STN6"
                     /translation="MPKRKKQNHHQPPTQQQPPLPEREETGDEEDGSPIAPNWIPPTC
                     PSSVEWINKLLCMHAMGLHCNENE"
     gene            join(71568. .71735,77584. .77630,77993. .78088,78248. .78377,
                     78523. .78560,78645. .78752,80106. .80180,80581. .80694,
                     82569. .82682,83241. .83363,83783. .83883,83980. .84065,
                     84367. .84506,84593. .84741,85435. .85518,85733. .85857,
                     86054. .86166,86309. .86480,86649. .86723,86816. .86886,
                     89863. .89895,89981. .90149,90324. .90388,90481. .90553,
                     90710. .>91487)
                     /gene="ABCF1"
                     /locus_tag="XXbac-BPG249D20.4-004"
     gene            join(71568. .71735,77584. .77630,77993. .78088,78251. .78377,
                     78523. .78560,78645. .78752,80106. .80180,80581. .80694,
                     82569. .82682,83241. .83363,83783. .83883,83980. .84065,
                     84367. .84506,84593. .84741,85435. .85518,85733. .85857,
                     86054. .86166,86309. .86480,86649. .86723,86816. .86886,
                     89863. .89895,89981. .90149,90324. .90388,90481. .90553,
                     90710. .>91487)
                     /gene="ABCF1"
                     /locus_tag="XXbac-BPG249D20.4-002"
     gene            join(71568. .71735,77584. .77630,77993. .78088,78251. .78377,
                     78523. .78560,78645. .78752,80106. .80180,80581. .80694,
                     83241. .83363,83783. .83883,83980. .84065,84367. .84506,
                     84593. .84741,85435. .85518,85733. .85857,86054. .86166,
                     86309. .86480,86649. .86723,86816. .86886,89863. .89895,
                     89981. .90149,90324. .90388,90481. .90553,90710. .91487)
                     /gene="ABCF1"
                     /locus_tag="XXbac-BPG249D20.4-001"
     mRNA            join(71568. .71735,77584. .77630,77993. .78088,78248. .78377,
                     78523. .78560,78645. .78752,80106. .80180,80581. .80694,

  Query Match             1.7%;  Score 52;  DB 8;  Length 151228;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12756 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12807


RESULT 125
AC040905/c
LOCUS       AC040905              151970 bp    DNA     linear   HTG 04-MAY-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-177A12 map 1, WORKING DRAFT
            SEQUENCE, 23 unordered pieces.
ACCESSION   AC040905
VERSION     AC040905.2  GI:7705170
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 151970)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 1, clone RP11-177A12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 151970)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 4, 2000 this sequence version replaced gi:7534048.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7581
                Center clone name: 177_A_12
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 139059 bases at least Q40
                Consensus quality: 145478 bases at least Q30
                Consensus quality: 148027 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 149770; sum-of-contigs
                Quality coverage: 3.7 in Q20 bases; agarose-fp
                Quality coverage: 4.1 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 23 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1023: contig of 1023 bp in length
            *     1024     1123: gap of 100 bp
            *     1124     2231: contig of 1108 bp in length
            *     2232     2331: gap of 100 bp
            *     2332     4468: contig of 2137 bp in length
            *     4469     4568: gap of 100 bp
            *     4569     5885: contig of 1317 bp in length
            *     5886     5985: gap of 100 bp
            *     5986     8879: contig of 2894 bp in length
            *     8880     8979: gap of 100 bp
            *     8980    10702: contig of 1723 bp in length
            *    10703    10802: gap of 100 bp
            *    10803    12995: contig of 2193 bp in length
            *    12996    13095: gap of 100 bp
            *    13096    15614: contig of 2519 bp in length
            *    15615    15714: gap of 100 bp
            *    15715    19162: contig of 3448 bp in length
            *    19163    19262: gap of 100 bp
            *    19263    23500: contig of 4238 bp in length
```

```
            *    23501    23600: gap of 100 bp
            *    23601    27529: contig of 3929 bp in length
            *    27530    27629: gap of 100 bp
            *    27630    32861: contig of 5232 bp in length
            *    32862    32961: gap of 100 bp
            *    32962    37601: contig of 4640 bp in length
            *    37602    37701: gap of 100 bp
            *    37702    44853: contig of 7152 bp in length
            *    44854    44953: gap of 100 bp
            *    44954    50843: contig of 5890 bp in length
            *    50844    50943: gap of 100 bp
            *    50944    60339: contig of 9396 bp in length
            *    60340    60439: gap of 100 bp
            *    60440    68661: contig of 8222 bp in length
            *    68662    68761: gap of 100 bp
            *    68762    78368: contig of 9607 bp in length
            *    78369    78468: gap of 100 bp
            *    78469    87095: contig of 8627 bp in length
            *    87096    87195: gap of 100 bp
            *    87196    97272: contig of 10077 bp in length
            *    97273    97372: gap of 100 bp
            *    97373   112327: contig of 14955 bp in length
            *   112328   112427: gap of 100 bp
            *   112428   133342: contig of 20915 bp in length
            *   133343   133442: gap of 100 bp
            *   133443   151970: contig of 18528 bp in length.
FEATURES             Location/Qualifiers
     source          1. .151970
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1"
                     /clone="RP11-177A12"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1023
                     /note="assembly_fragment"
     gap             1024. .1123
                     /estimated_length=100
     misc_feature    1124. .2231
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             2232. .2331
                     /estimated_length=100
     misc_feature    2332. .4468
                     /note="assembly_fragment"
     gap             4469. .4568
                     /estimated_length=100
     misc_feature    4569. .5885
                     /note="assembly_fragment"
     gap             5886. .5985
                     /estimated_length=100
     misc_feature    5986. .8879
                     /note="assembly_fragment"
     gap             8880. .8979
                     /estimated_length=100
     misc_feature    8980. .10702
                     /note="assembly_fragment"
     gap             10703. .10802
                     /estimated_length=100
     misc_feature    10803. .12995
                     /note="assembly_fragment"
     gap             12996. .13095
                     /estimated_length=100
     misc_feature    13096. .15614
                     /note="assembly_fragment"
     gap             15615. .15714
                     /estimated_length=100
     misc_feature    15715. .19162
                     /note="assembly_fragment"
     gap             19163. .19262
                     /estimated_length=100
     misc_feature    19263. .23500
                     /note="assembly_fragment"
     gap             23501. .23600
                     /estimated_length=100
     misc_feature    23601. .27529
                     /note="assembly_fragment"
     gap             27530. .27629
                     /estimated_length=100
     misc_feature    27630. .32861
                     /note="assembly_fragment"
     gap             32862. .32961
                     /estimated_length=100
     misc_feature    32962. .37601
                     /note="assembly_fragment"
     gap             37602. .37701
                     /estimated_length=100
     misc_feature    37702. .44853
                     /note="assembly_fragment"
     gap             44854. .44953
                     /estimated_length=100
     misc_feature    44954. .50843
                     /note="assembly_fragment"
     gap             50844. .50943
                     /estimated_length=100
     misc_feature    50944. .60339
                     /note="assembly_fragment"
     gap             60340. .60439
                     /estimated_length=100
     misc_feature    60440. .68661
                     /note="assembly_fragment"
     gap             68662. .68761
                     /estimated_length=100
     misc_feature    68762. .78368
                     /note="assembly_fragment"
     gap             78369. .78468
                     /estimated_length=100

  Query Match             1.7%;  Score 52;  DB 14;  Length 151970;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     100972 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 100921


RESULT 126
AC139547/c
LOCUS       AC139547              152040 bp    DNA     linear   HTG 21-FEB-2003
DEFINITION  Homo sapiens chromosome 15 clone RP13-425K1 map 15, WORKING DRAFT
            SEQUENCE, 24 unordered pieces.
ACCESSION   AC139547
VERSION     AC139547.2  GI:28460953
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152040)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP13-425K1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152040)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
```

```
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 152040)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 21, 2003 this sequence version replaced gi:28209587.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L29455
                Center clone name: 425_K_1
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 143383 bases at least Q40
                Consensus quality: 147021 bases at least Q30
                Consensus quality: 148441 bases at least Q20
                Insert size: 213000; agarose-fp
                Insert size: 149740; sum-of-contigs
                Quality coverage: 16.9 in Q20 bases; agarose-fp
                Quality coverage: 24.1 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 24 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      210: contig of 210 bp in length
            *     211      310: gap of 100 bp
            *     311     1698: contig of 1388 bp in length
            *    1699     1798: gap of 100 bp
            *    1799     2807: contig of 1009 bp in length
            *    2808     2907: gap of 100 bp
            *    2908     3937: contig of 1030 bp in length
            *    3938     4037: gap of 100 bp
            *    4038     4829: contig of 792 bp in length
            *    4830     4929: gap of 100 bp
            *    4930     5947: contig of 1018 bp in length
            *    5948     6047: gap of 100 bp
            *    6048     7740: contig of 1693 bp in length
            *    7741     7840: gap of 100 bp
            *    7841     9557: contig of 1717 bp in length
            *    9558     9657: gap of 100 bp
            *    9658    10567: contig of 910 bp in length
            *   10568    10667: gap of 100 bp
            *   10668    12556: contig of 1889 bp in length
            *   12557    12656: gap of 100 bp
            *   12657    14292: contig of 1636 bp in length
            *   14293    14392: gap of 100 bp
            *   14393    16890: contig of 2498 bp in length
            *   16891    16990: gap of 100 bp
            *   16991    17263: contig of 273 bp in length
            *   17264    17363: gap of 100 bp
            *   17364    19880: contig of 2517 bp in length
            *   19881    19980: gap of 100 bp
            *   19981    27529: contig of 7549 bp in length
            *   27530    27629: gap of 100 bp
            *   27630    29571: contig of 1942 bp in length
            *   29572    29671: gap of 100 bp
            *   29672    38820: contig of 9149 bp in length
            *   38821    38920: gap of 100 bp
            *   38921    46901: contig of 7981 bp in length
            *   46902    47001: gap of 100 bp
            *   47002    74440: contig of 27439 bp in length
            *   74441    74540: gap of 100 bp
            *   74541    85070: contig of 10530 bp in length
            *   85071    85170: gap of 100 bp
            *   85171    99159: contig of 13989 bp in length
            *   99160    99259: gap of 100 bp
            *   99260   118153: contig of 18894 bp in length
            *  118154   118253: gap of 100 bp
            *  118254   144364: contig of 26111 bp in length
            *  144365   144464: gap of 100 bp
            *  144465   152040: contig of 7576 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152040
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP13-425K1"
                     /clone_lib="RPCI-13 Human Female BAC"
     misc_feature    1. .210
                     /note="assembly_fragment"
     gap             211. .310
                     /estimated_length=100
     misc_feature    311. .1698
                     /note="assembly_fragment"
     gap             1699. .1798
                     /estimated_length=100
     misc_feature    1799. .2807
                     /note="assembly_fragment"
     gap             2808. .2907
                     /estimated_length=100
     misc_feature    2908. .3937
                     /note="assembly_fragment"
     gap             3938. .4037
                     /estimated_length=100
     misc_feature    4038. .4829
                     /note="assembly_fragment"
     gap             4830. .4929
                     /estimated_length=100
```

```
     misc_feature    4930. .5947
                     /note="assembly_fragment"
     gap             5948. .6047
                     /estimated_length=100
     misc_feature    6048. .7740
                     /note="assembly_fragment"
     gap             7741. .7840
                     /estimated_length=100
     misc_feature    7841. .9557
                     /note="assembly_fragment"
     gap             9558. .9657
                     /estimated_length=100
     misc_feature    9658. .10567
                     /note="assembly_fragment"
     gap             10568. .10667
                     /estimated_length=100
     misc_feature    10668. .12556
                     /note="assembly_fragment"
     gap             12557. .12656
                     /estimated_length=100
     misc_feature    12657. .14292
                     /note="assembly_fragment"
     gap             14293. .14392
                     /estimated_length=100
     misc_feature    14393. .16890
                     /note="assembly_fragment"
     gap             16891. .16990
                     /estimated_length=100
     misc_feature    16991. .17263

  Query Match             1.7%;  Score 52;  DB 14;  Length 152040;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74741 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 74690


RESULT 127
AC079863
LOCUS       AC079863               152492 bp    DNA     linear   PRI 21-JUN-2002
DEFINITION  Homo sapiens 3 BAC RP11-95L3 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC079863
VERSION     AC079863.11  GI:17105254
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152492)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 152492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 152492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 152492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 152492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 27, 2001 this sequence version replaced gi:16874814.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
```

continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

```
FEATURES             Location/Qualifiers
     source          1. .152492
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-95L3"
     misc_feature    1. .2000
                     /note="overlaps bases 166235. .168234 of clone AC092952"
                     /function="clone overlap"
     repeat_region   1562. .1757
                     /rpt_family="MIR"
     repeat_region   complement(2551. .2690)
                     /rpt_family="L1PA13"
     repeat_region   2920. .3079
                     /rpt_family="L2"
     repeat_region   3321. .3367
                     /rpt_family="MADE1"
     repeat_region   4166. .4312
                     /rpt_family="GA-rich"
     repeat_region   complement(4369. .4617)
                     /rpt_family="MIR"
     repeat_region   complement(4924. .5221)
                     /rpt_family="AluJo"
     repeat_region   complement(5853. .6280)
                     /rpt_family="L1"
     repeat_region   6281. .6339
                     /rpt_family="Tigger2"
     repeat_region   6346. .6785
                     /rpt_family="Tigger2"
     repeat_region   complement(6813. .6882)
                     /rpt_family="FAM"
     repeat_region   6924. .7222
                     /rpt_family="L1M3e"
     repeat_region   complement(7239. .7419)
                     /rpt_family="7SLRNA"
     repeat_region   7420. .7622
                     /rpt_family="Tigger1"
     repeat_region   complement(7627. .7975)
                     /rpt_family="L1M2"
     repeat_region   complement(7990. .8633)
                     /rpt_family="L1M4"
     repeat_region   8637. .8665
                     /rpt_family="AT_rich"
     repeat_region   8672. .8692
                     /rpt_family="AT_rich"
     repeat_region   8730. .8820
                     /rpt_family="MLT1E2"
     repeat_region   8833. .8861
                     /rpt_family="AT_rich"
     repeat_region   complement(8862. .8974)
                     /rpt_family="AluSq/x"
     repeat_region   9078. .9169
                     /rpt_family="MLT1E2"
     repeat_region   9272. .9486
                     /rpt_family="MLT1E2"
     repeat_region   complement(9489. .10203)
                     /rpt_family="L1M4"
     repeat_region   10423. .10584
                     /rpt_family="L1MCc"
     repeat_region   10891. .11733
                     /rpt_family="L1MCc"
     repeat_region   11832. .12106
                     /rpt_family="AluJo"
     repeat_region   12207. .12566
                     /rpt_family="L1MEc"
     repeat_region   complement(12623. .12990)
                     /rpt_family="MLT1D"
     repeat_region   complement(13033. .14014)
                     /rpt_family="L1M4"
     repeat_region   complement(14015. .14326)
                     /rpt_family="AluSc"
     repeat_region   complement(14327. .14539)
                     /rpt_family="L1M4"
     repeat_region   14540. .14823
                     /rpt_family="Tigger3b"
     repeat_region   complement(14824. .15113)
```

```
  Query Match             1.7%;  Score 52;  DB 8;  Length 152492;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151355 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 151406


RESULT 128
AC064817/c
LOCUS       AC064817              157135 bp    DNA     linear   HTG 16-JUL-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-46I22, WORKING DRAFT
            SEQUENCE, 27 unordered pieces.
ACCESSION   AC064817
VERSION     AC064817.4  GI:8980939
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 157135)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157135)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 9, 2000 this sequence version replaced gi:7770662.

            ---------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0046I22
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 141383 bases at least Q40
            Consensus quality: 146787 bases at least Q30
            Consensus quality: 149152 bases at least Q20
            Insert size: 165000; agarose-fp
```

```
            Insert size: 154535; sum-of-contigs
            Quality coverage: 3.28 in Q20 bases; agarose-fp
            Quality coverage: 3.81 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 27 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1120: contig of 1120 bp in length
            *    1121     1220: gap of unknown length
            *    1221     2618: contig of 1398 bp in length
            *    2619     2718: gap of unknown length
            *    2719     4808: contig of 2090 bp in length
            *    4809     4908: gap of unknown length
            *    4909     6698: contig of 1790 bp in length
            *    6699     6798: gap of unknown length
            *    6799     9212: contig of 2414 bp in length
            *    9213     9312: gap of unknown length
            *    9313    11332: contig of 2020 bp in length
            *   11333    11432: gap of unknown length
            *   11433    14518: contig of 3086 bp in length
            *   14519    14618: gap of unknown length
            *   14619    18220: contig of 3602 bp in length
            *   18221    18320: gap of unknown length
            *   18321    21383: contig of 3063 bp in length
            *   21384    21483: gap of unknown length
            *   21484    25924: contig of 4441 bp in length
            *   25925    26024: gap of unknown length
            *   26025    29223: contig of 3199 bp in length
            *   29224    29323: gap of unknown length
            *   29324    33796: contig of 4473 bp in length
            *   33797    33896: gap of unknown length
            *   33897    37526: contig of 3630 bp in length
            *   37527    37626: gap of unknown length
            *   37627    42125: contig of 4499 bp in length
            *   42126    42225: gap of unknown length
            *   42226    47482: contig of 5257 bp in length
            *   47483    47582: gap of unknown length
            *   47583    53156: contig of 5574 bp in length
            *   53157    53256: gap of unknown length
            *   53257    59938: contig of 6682 bp in length
            *   59939    60038: gap of unknown length
            *   60039    66057: contig of 6019 bp in length
            *   66058    66157: gap of unknown length
            *   66158    74668: contig of 8511 bp in length
            *   74669    74768: gap of unknown length
            *   74769    82448: contig of 7680 bp in length
            *   82449    82548: gap of unknown length
            *   82549    91047: contig of 8499 bp in length
            *   91048    91147: gap of unknown length
            *   91148    99649: contig of 8502 bp in length
            *   99650    99749: gap of unknown length
            *   99750   109752: contig of 10003 bp in length
            *  109753   109852: gap of unknown length
            *  109853   120015: contig of 10163 bp in length
            *  120016   120115: gap of unknown length
            *  120116   131890: contig of 11775 bp in length
            *  131891   131990: gap of unknown length
            *  131991   141546: contig of 9556 bp in length
            *  141547   141646: gap of unknown length
            *  141647   157135: contig of 15489 bp in length.
FEATURES             Location/Qualifiers
     source          1. .157135
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="RP11-46I22"
     misc_feature    1. .1120
                     /note="assembly_name:Contig15"
     gap             1121. .1220
                     /estimated_length=unknown
     misc_feature    1221. .2618
                     /note="assembly_name:Contig16
                     clone_end:SP6
                     vector_side:left"
     gap             2619. .2718
                     /estimated_length=unknown
     misc_feature    2719. .4808
                     /note="assembly_name:Contig17"
     gap             4809. .4908
                     /estimated_length=unknown
     misc_feature    4909. .6698
                     /note="assembly_name:Contig18"
     gap             6699. .6798
                     /estimated_length=unknown
     misc_feature    6799. .9212
                     /note="assembly_name:Contig19"
     gap             9213. .9312
                     /estimated_length=unknown
     misc_feature    9313. .11332
                     /note="assembly_name:Contig20"
     gap             11333. .11432
                     /estimated_length=unknown
     misc_feature    11433. .14518
                     /note="assembly_name:Contig21"
     gap             14519. .14618
                     /estimated_length=unknown
     misc_feature    14619. .18220
                     /note="assembly_name:Contig22"
     gap             18221. .18320
                     /estimated_length=unknown
     misc_feature    18321. .21383
                     /note="assembly_name:Contig23"
     gap             21384. .21483
                     /estimated_length=unknown
     misc_feature    21484. .25924
                     /note="assembly_name:Contig24"
     gap             25925. .26024
                     /estimated_length=unknown
     misc_feature    26025. .29223
                     /note="assembly_name:Contig25"
     gap             29224. .29323
                     /estimated_length=unknown
     misc_feature    29324. .33796
                     /note="assembly_name:Contig26"
     gap             33797. .33896
                     /estimated_length=unknown
     misc_feature    33897. .37526
                     /note="assembly_name:Contig27"
     gap             37527. .37626
                     /estimated_length=unknown
     misc_feature    37627. .42125
                     /note="assembly_name:Contig28"
     gap             42126. .42225
                     /estimated_length=unknown
     misc_feature    42226. .47482
                     /note="assembly_name:Contig29"
     gap             47483. .47582
                     /estimated_length=unknown
     misc_feature    47583. .53156
                     /note="assembly_name:Contig30"
     gap             53157. .53256
                     /estimated_length=unknown
     misc_feature    53257. .59938
                     /note="assembly_name:Contig31"
     gap             59939. .60038
                     /estimated_length=unknown
     misc_feature    60039. .66057
                     /note="assembly_name:Contig32"
     gap             66058. .66157
                     /estimated_length=unknown
```

```
     misc_feature    66158. .74668
                     /note="assembly_name:Contig33"
     gap             74669. .74768
                     /estimated_length=unknown
     misc_feature    74769. .82448
                     /note="assembly_name:Contig34"
     gap             82449. .82548
                     /estimated_length=unknown
     misc_feature    82549. .91047
                     /note="assembly_name:Contig35"
     gap             91048. .91147
                     /estimated_length=unknown
     misc_feature    91148. .99649
                     /note="assembly_name:Contig36"
     gap             99650. .99749

  Query Match             1.7%;  Score 52;  DB 14;  Length 157135;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147583 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 147532


RESULT 129
AC008677
LOCUS       AC008677               160929 bp    DNA     linear   PRI 01-JAN-2002
DEFINITION  Homo sapiens chromosome 5 clone CTB-47B8, complete sequence.
ACCESSION   AC008677
VERSION     AC008677.5  GI:18030123
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 160929)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160929)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 160929)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 1, 2002 this sequence version replaced gi:7709258.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .160929
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-47B8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 160929;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

-Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      92839 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 92890


RESULT 130
AC027706/c
LOCUS       AC027706               161970 bp    DNA     linear   HTG 02-MAY-2000
DEFINITION  Homo sapiens chromosome 12 clone RP11-665C13 map 12, WORKING DRAFT
            SEQUENCE, 49 unordered pieces.
ACCESSION   AC027706
VERSION     AC027706.2  GI:7677821
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 161970)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 12, clone RP11-665C13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 161970)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,P.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 2, 2000 this sequence version replaced gi:7382510.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6713
                Center clone name: 665_C_13
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 140943 bases at least Q40
                Consensus quality: 150450 bases at least Q30
                Consensus quality: 154599 bases at least Q20
                Insert size: 165000; agarose-fp
                Insert size: 157170; sum-of-contigs
                Quality coverage: 3.1 in Q20 bases; agarose-fp
                Quality coverage: 3.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 49 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
```

```
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1052: contig of 1052 bp in length
*     1053     1152: gap of 100 bp
*     1153     2378: contig of 1226 bp in length
*     2379     2478: gap of 100 bp
*     2479     3812: contig of 1334 bp in length
*     3813     3912: gap of 100 bp
*     3913     5368: contig of 1456 bp in length
*     5369     5468: gap of 100 bp
*     5469     6519: contig of 1051 bp in length
*     6520     6619: gap of 100 bp
*     6620     7852: contig of 1233 bp in length
*     7853     7952: gap of 100 bp
*     7953     8464: contig of 512 bp in length
*     8465     8564: gap of 100 bp
*     8565    10347: contig of 1783 bp in length
*    10348    10447: gap of 100 bp
*    10448    12145: contig of 1698 bp in length
*    12146    12245: gap of 100 bp
*    12246    14006: contig of 1761 bp in length
*    14007    14106: gap of 100 bp
*    14107    15358: contig of 1252 bp in length
*    15359    15458: gap of 100 bp
*    15459    17152: contig of 1694 bp in length
*    17153    17252: gap of 100 bp
*    17253    18800: contig of 1548 bp in length
*    18801    18900: gap of 100 bp
*    18901    20291: contig of 1391 bp in length
*    20292    20391: gap of 100 bp
*    20392    21889: contig of 1498 bp in length
*    21890    21989: gap of 100 bp
*    21990    24355: contig of 2366 bp in length
*    24356    24455: gap of 100 bp
*    24456    26757: contig of 2302 bp in length
*    26758    26857: gap of 100 bp
*    26858    28317: contig of 1460 bp in length
*    28318    28417: gap of 100 bp
*    28418    30781: contig of 2364 bp in length
*    30782    30881: gap of 100 bp
*    30882    32789: contig of 1908 bp in length
*    32790    32889: gap of 100 bp
*    32890    35214: contig of 2325 bp in length
*    35215    35314: gap of 100 bp
*    35315    37181: contig of 1867 bp in length
*    37182    37281: gap of 100 bp
*    37282    39642: contig of 2361 bp in length
*    39643    39742: gap of 100 bp
*    39743    41978: contig of 2236 bp in length
*    41979    42078: gap of 100 bp
*    42079    44963: contig of 2885 bp in length
*    44964    45063: gap of 100 bp
*    45064    47265: contig of 2202 bp in length
*    47266    47365: gap of 100 bp
*    47366    49847: contig of 2482 bp in length
*    49848    49947: gap of 100 bp
*    49948    52474: contig of 2527 bp in length
*    52475    52574: gap of 100 bp
*    52575    54767: contig of 2193 bp in length
*    54768    54867: gap of 100 bp
*    54868    57938: contig of 3071 bp in length
*    57939    58038: gap of 100 bp
*    58039    60538: contig of 2500 bp in length
*    60539    60638: gap of 100 bp
*    60639    63422: contig of 2784 bp in length
*    63423    63522: gap of 100 bp
*    63523    66038: contig of 2516 bp in length
*    66039    66138: gap of 100 bp
*    66139    69185: contig of 3047 bp in length
*    69186    69285: gap of 100 bp
*    69286    72321: contig of 3036 bp in length
*    72322    72421: gap of 100 bp
*    72422    75240: contig of 2819 bp in length
*    75241    75340: gap of 100 bp
*    75341    77598: contig of 2258 bp in length
*    77599    77698: gap of 100 bp
*    77699    82485: contig of 4787 bp in length
*    82486    82585: gap of 100 bp
*    82586    86177: contig of 3592 bp in length
*    86178    86277: gap of 100 bp
*    86278    91129: contig of 4852 bp in length
*    91130    91229: gap of 100 bp
*    91230    94475: contig of 3246 bp in length
*    94476    94575: gap of 100 bp
*    94576    98382: contig of 3807 bp in length
*    98383    98482: gap of 100 bp
*    98483   103891: contig of 5409 bp in length
*   103892   103991: gap of 100 bp
*   103992   109749: contig of 5758 bp in length
*   109750   109849: gap of 100 bp
*   109850   116075: contig of 6226 bp in length
*   116076   116175: gap of 100 bp
*   116176   122437: contig of 6262 bp in length
*   122438   122537: gap of 100 bp
*   122538   132392: contig of 9855 bp in length
*   132393   132492: gap of 100 bp
*   132493   141881: contig of 9389 bp in length
*   141882   141981: gap of 100 bp
*   141982   161970: contig of 19989 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161970
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12"
                     /clone="RP11-665C13"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1052
                     /note="assembly_fragment"
     gap             1053. .1152
                     /estimated_length=100
     misc_feature    1153. .2378
                     /note="assembly_fragment"
     gap             2379. .2478
                     /estimated_length=100
     misc_feature    2479. .3812
                     /note="assembly_fragment"
     gap             3813. .3912
                     /estimated_length=100
     misc_feature    3913. .5368
                     /note="assembly_fragment"
     gap             5369. .5468
                     /estimated_length=100
     misc_feature    5469. .6519
                     /note="assembly_fragment"
     gap             6520. .6619
                     /estimated_length=100
     misc_feature    6620. .7852
                     /note="assembly_fragment"

  Query Match             1.7%;  Score 52;  DB 14;  Length 161970;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5801 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 5750


RESULT 131
AL662822
LOCUS       AL662822              162948 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-17K10 on chromosome 6 contains
```

```
          a major histocompatibility complex class 1-related polypeptide
          sequence B (MICB) pseudogene, a  thymopoietin (TMPO) pseudogene, a
          succinate-CoA ligase, ADP forming, beta subunit (SUCLA2)
          pseudogene, a RAN, member RAS oncogene family pseudogene, the HLA-E
          gene for major histocompatibility complex, class I, E and six CpG
          islands, complete sequence.
ACCESSION AL662822
VERSION   AL662822.4  GI:18650691
KEYWORDS  HTG; A-BETA; ARA24; CpG island; HLA-6.2; HLA-E; major
          histocompatibility complex class 1-related p; major
          histocompatibility complex, class I, E; MICB; RAN; succinate-CoA
          ligase; SUCLA2; TC4; thymopoietin; TMPO; TP.
SOURCE    Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1  (bases 1 to 162948)
  AUTHORS Johnson,C.
  TITLE   Direct Submission
  JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
          Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Feb 11, 2002 this sequence version replaced gi:18135278.
          The following abbreviations are used to associate primary accession
          numbers given in the feature table with their source databases:
          Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
          on the WORMPEP database can be found at
          http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-17K10 is
          from a CHORI-502 human bac - COX cell line library VECTOR:
          pTARBAC2.1
          This sequence was generated from part of bacterial clone contigs
          constructed by the MHC Haplotype Consortium and collaborators.
          Further information can be found at
          http://www.sanger.ac.uk/HGP/Chr6/MHC
          -------------- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: vega@sanger.ac.uk
          --------------
          This sequence was finished as follows unless otherwise noted: all
          regions were either double-stranded or sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by at least
          one subclone; and the assembly was confirmed by restriction digest,
          except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .162948
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-17K10"
                     /clone_lib="CHORI-502"
     misc_feature    1
                     /note="Clone_left_end: XXbac-17K10"
     misc_feature    28257
                     /note="Clone_right_end: XXbac-22J1"
     gene            join(32973. .33223,33490. .33775,34377. .34894,37463. .37598)
                     /locus_tag="XXbac-BCX17K10.1-001"
                     /pseudo
     CDS             join(32973. .33223,33490. .33775,34377. .34894,37463. .37598)
                     /locus_tag="XXbac-BCX17K10.1-001"
                     /note="match: proteins: O98234 O98266 O98269 Q8WLQ0 Q8WLQ1
                     Q8WLZ6 Q95HA2 Q9BCR4 Q9BCR5 Q9GIZ5 Q9TPU8"
                     /pseudo
                     /codon_start=1
                     /product="MHC class I polypeptide-related sequence C"
     misc_feature    79963
                     /note="Clone_left_end: XXbac-98J21"
     gene            join(84827. .85246,85498. .86253)
                     /locus_tag="XXbac-BCX17K10.3-001"
                     /pseudo
     CDS             join(84827. .85246,85498. .86253)
                     /locus_tag="XXbac-BCX17K10.3-001"
                     /note="match: proteins: AAC25390 O73766 P42167 Q61028
                     Q61030 Q61031 Q61032"
                     /pseudo
                     /codon_start=1
                     /product="thymopoietin-like 1"
     gene            87207. .88544
                     /locus_tag="XXbac-BCX17K10.5-001"
                     /pseudo
     CDS             87207. .88544
                     /locus_tag="XXbac-BCX17K10.5-001"
                     /note="match: proteins: O97580 P53312 P53588 Q8WM98 Q9P2R7
                     Q9Y4T0 Q9YI37 Q9Z2I9"
                     /pseudo
                     /codon_start=1
                     /product="succinate-CoA ligase, ADP-forming, beta subunit
                     pseudogene"
     gene            104251. .104901
                     /locus_tag="XXbac-BCX17K10.2-001"
                     /pseudo
     CDS             104251. .104901
                     /locus_tag="XXbac-BCX17K10.2-001"
                     /note="match: proteins: AAH14518 AAH14829 AAM33416 O76173
                     P17080 P32835 P35292 P42558 P91857 Q22908 Q61820 Q9IBE8
                     Q9XJ45 Q9YGC0"
                     /pseudo
                     /codon_start=1
                     /product="RAN, member RAS oncogene family pseudogene 1"
     gene            107752. .112507
                     /gene="HLA-E"
                     /locus_tag="XXbac-BCX17K10.4-001"
     mRNA            join(107752. .107905,108036. .108305,108550. .108825,
                     109447. .109722,109847. .109963,110714. .110746,
                     110851. .110893,111059. .112507)
                     /gene="HLA-E"
                     /locus_tag="XXbac-BCX17K10.4-001"
                     /product="major histocompatibility complex, class I, E"
                     /note="match: ESTs: BG028524.1 BG756822.1 BM791769.1
                     BM917068.1
                     match: cDNAs: AJ293263.1 AJ293264.1 BC002578.1 M20022.1"
     CDS             join(107842. .107905,108036. .108305,108550. .108825,
                     109447. .109722,109847. .109963,110714. .110746,
                     110851. .110891)
                     /gene="HLA-E"
                     /locus_tag="XXbac-BCX17K10.4-001"
                     /standard_name="OTTHUMP00000029626"
                     /note="match: proteins: Q9GIY7 Q9GIY8"
                     /codon_start=1
                     /product="major histocompatibility complex, class I, E"
                     /protein_id="CAI17548.1"
                     /db_xref="GI:55961319"
                     /db_xref="InterPro:IPR001039"
                     /db_xref="InterPro:IPR003006"
                     /db_xref="InterPro:IPR003597"
                     /db_xref="InterPro:IPR007110"
                     /translation="MVDGTLLLLLSEALALTQTWAGSHSLKYFHTSVSRPGRGEPRFI
                     SVGYVDDTQFVRFDNDAASPRMVPRAPWMEQEGSEYWDRETRSARDTAQIFRVNLRTL
                     RGYYNQSEAGSHTLQWMHGCELGPDRRFLRGYEQFAYDGKDYLTLNEDLRSWTAVDTA
                     AQISEQKSNDASEAEHQRAYLEDTCVEWLHKYLEKGKETLLHLEPPKTHVTHHPISDH
                     EATLRCWALGFYPAEITLTWQQDGEGHTQDTELVETRPAGDGTFQKWAAVVVPSGEEQ
                     RYTCHVQHEGLPEPVTLRWKPASQPTIPIVGIIAGLVLLGSVVSGAVVAAVIWRKKSS
                     GGKGGSYSKAEWSDSAQGSESHSL"
     misc_feature    162948
                     /note="Clone_right_end: XXbac-17K10"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 162948;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130886 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 130937


RESULT 132
AC108670/c
LOCUS       AC108670               163157 bp    DNA     linear   PRI 28-DEC-2002
DEFINITION  Homo sapiens 3 BAC RP11-218A22 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC108670
VERSION     AC108670.10  GI:23355671
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 163157)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 163157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 163157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 163157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 163157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 30, 2002 this sequence version replaced gi:20514653.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
            ml.
FEATURES             Location/Qualifiers
     source          1. .163157
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-218A22"
     repeat_region   1. .234
                     /rpt_family="AluSg/x"
     repeat_region   348. .468
                     /rpt_family="FLAM_C"
     repeat_region   469. .765
                     /rpt_family="AluSx"
     repeat_region   1012. .1319
                     /rpt_family="AluSx"
     repeat_region   complement(1553. .1867)
```

```
                     /rpt_family="AluJb"
     repeat_region   complement(1868. .1977)
                     /rpt_family="AluJo"
     repeat_region   2381. .2686
                     /rpt_family="AluSx"
     repeat_region   2687. .2756
                     /rpt_family="(TA)n"
     repeat_region   complement(2757. .2829)
                     /rpt_family="Ricksha"
     repeat_region   complement(2830. .3124)
                     /rpt_family="AluSq"
     repeat_region   complement(3125. .3254)
                     /rpt_family="Ricksha"
     repeat_region   3262. .3847
                     /rpt_family="LTR9"
     repeat_region   3966. .4259
                     /rpt_family="AluSx"
     repeat_region   complement(4327. .4607)
                     /rpt_family="HERVL"
     repeat_region   4833. .5124
                     /rpt_family="AluSq"
     repeat_region   5125. .5436
                     /rpt_family="AluSx"
     repeat_region   5635. .5683
                     /rpt_family="AT_rich"
     repeat_region   complement(5746. .5934)
                     /rpt_family="AluJb"
     repeat_region   complement(5935. .6248)
                     /rpt_family="AluSx"
     repeat_region   complement(6249. .6350)
                     /rpt_family="AluJb"
     repeat_region   6613. .6747
                     /rpt_family="MER112"
     repeat_region   complement(6837. .7158)
                     /rpt_family="AluSc"
     repeat_region   7439. .7727
                     /rpt_family="AluSp"
     repeat_region   7738. .7771
                     /rpt_family="(TAAA)n"
     repeat_region   complement(7828. .7948)
                     /rpt_family="MLT1H"
     repeat_region   complement(8132. .8280)
                     /rpt_family="MLT1G1"
     repeat_region   complement(8357. .8467)
                     /rpt_family="L2"
     repeat_region   complement(8472. .8775)
                     /rpt_family="AluSg"
     repeat_region   complement(9131. .9257)
                     /rpt_family="L2"
     repeat_region   9246. .9369
                     /rpt_family="MIR"
     repeat_region   9730. .9751
                     /rpt_family="AT_rich"
     repeat_region   9761. .9893
                     /rpt_family="AluSq/x"
     repeat_region   9907. .10200
                     /rpt_family="AluSx"
     repeat_region   complement(10355. .10680)
                     /rpt_family="L2"
     repeat_region   11039. .11148
                     /rpt_family="MIR"

  Query Match             1.7%;  Score 52;  DB 8;  Length 163157;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22314 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 22263


RESULT 133
AC141308
```

```
LOCUS       AC141308              163280 bp    DNA     linear   HTG 11-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-855H8, WORKING DRAFT
            SEQUENCE, 3 unordered pieces.
ACCESSION   AC141308
VERSION     AC141308.1  GI:28913088
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 163280)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163280)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 1623000
            Center clone name: RPCI-11_855H8
            -------
            Summary Statistics
            Consensus quality: 162905 bases at least Q40
            Consensus quality: 162980 bases at least Q30
            Consensus quality: 163047 bases at least Q20
            Estimated insert size: 175000; agarose-fp estimation
            Estimated insert size: 163080; sum-of-contigs estimation
            Quality coverage:  13.59 in Q20 bases; agarose-fp estimation
            Quality coverage:  14.58 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    22836: contig of 22836 bp in length
            *    22837    22936: gap of unknown length
            *    22937    49828: contig of 26892 bp in length
            *    49829    49928: gap of unknown length
            *    49929   163280: contig of 113352 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163280
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-855H8"
                     /clone_lib="RPCI human BAC library 11"
     gap             22837. .22936
                     /estimated_length=unknown
     gap             49829. .49928
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 163280;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134135 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 134186
```

```
RESULT 134
AC161476
LOCUS       AC161476              163432 bp    DNA     linear   HTG 24-JUN-2005
DEFINITION  Pan troglodytes chromosome UNKNOWN clone CH251-284F7, WORKING DRAFT
            SEQUENCE, 16 unordered pieces.
ACCESSION   AC161476
VERSION     AC161476.1  GI:65146782
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 163432)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163432)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 163432)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUN-2005) Genetics, Genome Sequencing Center, 4444
.           Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
            ---------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: C_AB0284F07
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 152742 bases at least Q40
            Consensus quality: 156159 bases at least Q30
            Consensus quality: 158103 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1144: contig of 1144 bp in length
            *     1145     1244: gap of unknown length
            *     1245     2840: contig of 1596 bp in length
            *     2841     2940: gap of unknown length
            *     2941     5060: contig of 2120 bp in length
            *     5061     5160: gap of unknown length
            *     5161     8540: contig of 3380 bp in length
            *     8541     8640: gap of unknown length
            *     8641    11235: contig of 2595 bp in length
            *    11236    11335: gap of unknown length
            *    11336    18780: contig of 7445 bp in length
            *    18781    18880: gap of unknown length
            *    18881    24016: contig of 5136 bp in length
            *    24017    24116: gap of unknown length
            *    24117    30109: contig of 5993 bp in length
            *    30110    30209: gap of unknown length
            *    30210    42865: contig of 12656 bp in length
            *    42866    42965: gap of unknown length
            *    42966    51561: contig of 8596 bp in length
            *    51562    51661: gap of unknown length
            *    51662    65150: contig of 13489 bp in length
            *    65151    65250: gap of unknown length
            *    65251    74190: contig of 8940 bp in length
            *    74191    74290: gap of unknown length
            *    74291    84891: contig of 10601 bp in length
            *    84892    84991: gap of unknown length
            *    84992   103199: contig of 18208 bp in length
            *   103200   103299: gap of unknown length
            *   103300   125871: contig of 22572 bp in length
            *   125872   125971: gap of unknown length
            *   125972   163432: contig of 37461 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163432
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNKNOWN"
                     /clone="CH251-284F7"
     misc_feature    1. .1144
                     /note="assembly_name:Contig41"
     gap             1145. .1244
                     /estimated_length=unknown
     misc_feature    1245. .2840
                     /note="assembly_name:Contig45"
     gap             2841. .2940
                     /estimated_length=unknown
     misc_feature    2941. .5060
                     /note="assembly_name:Contig50"
     gap             5061. .5160
                     /estimated_length=unknown
     misc_feature    5161. .8540
                     /note="assembly_name:Contig51"
     gap             8541. .8640
                     /estimated_length=unknown
     misc_feature    8641. .11235
                     /note="assembly_name:Contig52"
     gap             11236. .11335
                     /estimated_length=unknown
     misc_feature    11336. .18780
                     /note="assembly_name:Contig53"
     gap             18781. .18880
                     /estimated_length=unknown
     misc_feature    18881. .24016
                     /note="assembly_name:Contig54"
     gap             24017. .24116
                     /estimated_length=unknown
     misc_feature    24117. .30109
                     /note="assembly_name:Contig55"
     gap             30110. .30209
                     /estimated_length=unknown
     misc_feature    30210. .42865
                     /note="assembly_name:Contig56"
     gap             42866. .42965
                     /estimated_length=unknown
     misc_feature    42966. .51561
                     /note="assembly_name:Contig57
                     clone_end:SP6
                     vector_side:right"
     gap             51562. .51661
                     /estimated_length=unknown
     misc_feature    51662. .65150
                     /note="assembly_name:Contig58"
     gap             65151. .65250
                     /estimated_length=unknown
     misc_feature    65251. .74190
                     /note="assembly_name:Contig59"
     gap             74191. .74290
                     /estimated_length=unknown
     misc_feature    74291. .84891
                     /note="assembly_name:Contig60"
     gap             84892. .84991
                     /estimated_length=unknown
```

```
     misc_feature    84992. .103199
                     /note="assembly_name:Contig61"
     gap             103200. .103299
                     /estimated_length=unknown
     misc_feature    103300. .125871
                     /note="assembly_name:Contig62"
     gap             125872. .125971
                     /estimated_length=unknown
     misc_feature    125972. .163432
                     /note="assembly_name:Contig63"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 163432;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91488 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 91539


RESULT 135
AC092606
LOCUS       AC092606               163437 bp    DNA     linear   PRI 08-OCT-2003
DEFINITION  Homo sapiens BAC clone RP11-150B23 from 7, complete sequence.
ACCESSION   AC092606 AC025297
VERSION     AC092606.2  GI:15638770
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 163437)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 163437)
  AUTHORS   Abbott,A., Doebber,A. and Hou,S.
  TITLE     The sequence of Homo sapiens BAC clone RP11-150B23
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 163437)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 163437)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 163437)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 163437)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 163437)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   8  (bases 1 to 163437)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Sep 18, 2001 this sequence version replaced gi:14916191.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0150B23
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            Data from AC092148 was used to finish this clone, AC092606.

            The clone sequenced to the right is RP11-95L23.  Actual start of
            this clone is at base position 1 of RP11-150B23 actual end is at
            base position 163437 of RP11-150B23.

            The sequence of AC025297 has been incorporated into AC092606.
FEATURES             Location/Qualifiers
     source          1. .163437
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP11-150B23"
                     /clone_lib="RPCI-11"
     repeat_region   1. .953
                     /rpt_family="L1"
     repeat_region   1041. .1227
                     /rpt_family="L1"
     repeat_region   1227. .1868
                     /rpt_family="L1"
     repeat_region   3283. .3321
                     /rpt_family="(TTTG)n"
     repeat_region   10219. .10601
                     /rpt_family="MaLR"
     repeat_region   10602. .12171
```

```
                     /rpt_family="MaLR"
     repeat_region   12172. .12542
                     /rpt_family="MaLR"
     repeat_region   13005. .13032
                     /rpt_family="AT_rich"
     repeat_region   13940. .14015
                     /rpt_family="MIR"
     repeat_region   14231. .14531
                     /rpt_family="Alu"
     repeat_region   16011. .16093
                     /rpt_family="MIR"
     repeat_region   17056. .17085
                     /rpt_family="AT_rich"
     repeat_region   18554. .18884
                     /rpt_family="Alu"
     repeat_region   19610. .19797
                     /rpt_family="MIR"
     repeat_region   19838. .19908
                     /rpt_family="L2"
     repeat_region   19909. .20298
                     /rpt_family="MaLR"
     repeat_region   21626. .21672
                     /rpt_family="L2"
     repeat_region   22177. .22416
                     /rpt_family="Alu"
     misc_feature    22180. .22419
                     /note="CpG_island (%GC=58.3, o/e=0.86, #CpGs=18)"
     repeat_region   22422. .22455
                     /rpt_family="(A)n"
     repeat_region   22747. .22869
                     /rpt_family="MER1_type"
     repeat_region   22871. .23099
                     /rpt_family="MaLR"
     repeat_region   23115. .23659
                     /rpt_family="ERVL"
     repeat_region   24116. .24359
                     /rpt_family="L2"
     repeat_region   24900. .24942
                     /rpt_family="(CATA)n"
     repeat_region   25007. .25236
                     /rpt_family="MIR"
     repeat_region   25878. .26146
                     /rpt_family="Alu"
     repeat_region   27029. .27127
                     /rpt_family="L2"
     repeat_region   27128. .27167
                     /rpt_family="(CA)n"
     repeat_region   27168. .27388
                     /rpt_family="L2"
     repeat_region   28936. .29055
                     /rpt_family="MIR"
     repeat_region   29766. .29797
                     /rpt_family="(TA)n"
     repeat_region   29878. .29904
                     /rpt_family="(GA)n"
     repeat_region   29906. .30195
                     /rpt_family="Alu"
     repeat_region   30620. .30777
                     /rpt_family="L1"
     repeat_region   30778. .31086
                     /rpt_family="Alu"
     repeat_region   31087. .31289
                     /rpt_family="L1"
     repeat_region   32124. .32189
                     /rpt_family="L2"
     repeat_region   32323. .32442
                     /rpt_family="CR1"
     repeat_region   33086. .33222
                     /rpt_family="MIR"
     repeat_region   33274. .33551
                     /rpt_family="Alu"
     repeat_region   33822. .34122
                     /rpt_family="MaLR"
     repeat_region   34435. .34711
                     /rpt_family="L2"
     repeat_region   35362. .35457
                     /rpt_family="CT-rich"

  Query Match             1.7%;  Score 52;  DB 8;  Length 163437;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85364 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 85415


RESULT 136
AC096508/c
LOCUS       AC096508              163662 bp    DNA     linear   HTG 13-SEP-2002
DEFINITION  Homo sapiens chromosome X clone RP11-111F17, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC096508
VERSION     AC096508.8  GI:22830224
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 163662)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163662)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 163662)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 13, 2002 this sequence version replaced gi:20376857.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Drafting Center Code: BCM
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: HDML
                Center clone name: RP11-111F17
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Sequencing vector: M13;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 161410 bases at least Q40
                Consensus quality: 162039 bases at least Q30
                Consensus quality: 162473 bases at least Q20
                Estimated insert size: 164107; sum-of-contigs estimation
                Quality coverage: 10x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3227: contig of 3227 bp in length
            *     3228     3327: gap of unknown length
            *     3328    12653: contig of 9326 bp in length
            *    12654    12753: gap of unknown length
            *    12754    33426: contig of 20673 bp in length
            *    33427    33526: gap of unknown length
            *    33527    60991: contig of 27465 bp in length
            *    60992    61091: gap of unknown length
            *    61092   163662: contig of 102571 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163662
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="RP11-111F17"
     gap             3228. .3327
                     /estimated_length=unknown
     gap             12654. .12753
                     /estimated_length=unknown
     gap             33427. .33526
                     /estimated_length=unknown
     gap             60992. .61091
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 163662;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85350 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85299
```

```
RESULT 137
AC150837/c
LOCUS       AC150837              164429 bp    DNA     linear   HTG 02-APR-2005
DEFINITION  Papio anubis clone RP41-147B6, WORKING DRAFT SEQUENCE, 16 ordered
            pieces.
ACCESSION   AC150837
VERSION     AC150837.3  GI:62175061
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 164429)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Fuksenko,T., Gestole,M., Greene,A., Guan,X., Gupta,J.,
            Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hunter,G.,
            Hurle,B., Idol,J.R., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q.,
            Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
            Masiello,C., Maskeri,B., McDowell,J., Mojidi,H.A., Mullikin,J.C.,
            Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A., Puri,O.,
            Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
            Stantripop,S., Stephen,E., Taye,A., Thomas,J.W., Thomas,P.J.,
            Tsipouri,V., Ung,L., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164429)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 164429)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2005) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Apr 2, 2005 this sequence version replaced gi:56159036.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: hql
                 Center clone name: 147B06

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is generally based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 160267 bases at least Q40
                 Consensus quality: 161522 bases at least Q30
                 Consensus quality: 162357 bases at least Q20
                 Insert size: 132000; agarose-fp
                 Insert size: 162929; sum-of-contigs
                 Quality coverage: 12.99x in Q20 bases; agarose-fp
                 Quality coverage: 10.53x in Q20 bases; sum-of-contigs
```

```
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     4345: contig of 4345 bp in length
            *     4346     4445: gap of unknown length
            *     4446    11449: contig of 7004 bp in length
            *    11450    11549: gap of unknown length
            *    11550    28453: contig of 16904 bp in length
            *    28454    28553: gap of unknown length
            *    28554    30919: contig of 2366 bp in length
            *    30920    31019: gap of unknown length
            *    31020    36197: contig of 5178 bp in length
            *    36198    36297: gap of unknown length
            *    36298    39674: contig of 3377 bp in length
            *    39675    39774: gap of unknown length
            *    39775    46430: contig of 6656 bp in length
            *    46431    46530: gap of unknown length
            *    46531    49597: contig of 3067 bp in length
            *    49598    49697: gap of unknown length
            *    49698    63075: contig of 13378 bp in length
            *    63076    63175: gap of unknown length
            *    63176    66462: contig of 3287 bp in length
            *    66463    66562: gap of unknown length
            *    66563    69322: contig of 2760 bp in length
            *    69323    69422: gap of unknown length
            *    69423    85463: contig of 16041 bp in length
            *    85464    85563: gap of unknown length
            *    85564   119974: contig of 34411 bp in length
            *   119975   120074: gap of unknown length
            *   120075   131535: contig of 11461 bp in length
            *   131536   131635: gap of unknown length
            *   131636   133931: contig of 2296 bp in length
            *   133932   134031: gap of unknown length
            *   134032   164429: contig of 30398 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164429
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="RP41-147B6"
                     /clone_lib="RP41"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .50701
                     /note="clone overlaps with GenBank Accession Number
                     AC094014 clone RP41-371B12 (center project name ccv)"
     misc_feature    1. .4345
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             4346. .4445
                     /estimated_length=unknown
     misc_feature    4446. .11449
                     /note="assembly_fragment"
     gap             11450. .11549
                     /estimated_length=unknown
     misc_feature    11550. .28453
                     /note="assembly_fragment"
     gap             28454. .28553
                     /estimated_length=unknown
     misc_feature    28554. .30919
                     /note="assembly_fragment"
     gap             30920. .31019
                     /estimated_length=unknown
     misc_feature    31020. .36197
                     /note="assembly_fragment"
     gap             36198. .36297
                     /estimated_length=unknown
     misc_feature    36298. .39674
                     /note="assembly_fragment"
     gap             39675. .39774
                     /estimated_length=unknown
     misc_feature    39775. .46430
                     /note="assembly_fragment"
     gap             46431. .46530
                     /estimated_length=unknown
     misc_feature    46531. .49597
                     /note="assembly_fragment"
     gap             49598. .49697
                     /estimated_length=unknown
     misc_feature    49698. .63075
                     /note="assembly_fragment"
     gap             63076. .63175
                     /estimated_length=unknown
     misc_feature    63176. .66462
                     /note="assembly_fragment"
     gap             66463. .66562
                     /estimated_length=unknown
     misc_feature    66563. .69322
                     /note="assembly_fragment"
     gap             69323. .69422
                     /estimated_length=unknown
     misc_feature    69423. .85463
                     /note="assembly_fragment"
     misc_feature    74551. .164429
                     /note="clone overlaps with GenBank Accession Number
                     AC093126 clone RP41-231J21 (center project name ccu)"
     gap             85464. .85563
                     /estimated_length=unknown
     misc_feature    85564. .119974
                     /note="assembly_fragment"
     gap             119975. .120074
                     /estimated_length=unknown
     misc_feature    120075. .131535
                     /note="assembly_fragment"
     gap             131536. .131635
                     /estimated_length=unknown
     misc_feature    131636. .133931
                     /note="assembly_fragment"
     gap             133932. .134031
                     /estimated_length=unknown
     misc_feature    134032. .164429
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 164429;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110576 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 110525


RESULT 138
AC097061/c
LOCUS       AC097061              165311 bp    DNA     linear   PRI 04-JAN-2002
DEFINITION  Homo sapiens chromosome 1 clone RP11-415L24, complete sequence.
ACCESSION   AC097061
VERSION     AC097061.2  GI:18057178
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 165311)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and
```

Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 165311)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 165311)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (04-JAN-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
COMMENT On Jan 4, 2002 this sequence version replaced gi:15982504.
---------- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchtgs@u.washington.edu
---------- Project Information
Center project name: chr-1
Center clone name: RP11-415L24 (sc0692)
---------- Summary Statistics
Sequencing vector: plasmid; 62% of reads
Sequencing vector: plasmid; L08752; 38% of reads
Chemistry: Dye-terminator ET; 92% of reads
Chemistry: Dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165135 bases at least Q40
Consensus quality: 165301 bases at least Q30
Consensus quality: 165311 bases at least Q20
Insert size: 165311; sum-of-contigs
Quality coverage: 10.5x in Q20 bases; sum-of-contigs
----------.
Overlapping Sequences:
5': Mapping in progress
3': Mapping in progress

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI		HindIII		BglII	
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
8696	8888	2946	2958	3938	3943
6	<800	6382	6573	2067	2063
1114	1095	512	<800	6223	5919
2678	2839	449	<800	3058	3059
12640	12720	1478	1474	1596	1675
3027	3001	4425	4355	1864	1854
3092	3205	4873	4812	2371	2419
3033	3001	4854	4812	2950	3177
2024	2026	101	<800	66	<800
5868	5803	6408	6573	1920	1854
2058	2026	917	912	8892	8587
5042	4888	650	<800	3816	3591
4032	4019	13	<800	1133	1206
3619	3528	1690	1726	4202	4243
591	<800	386	<800	3573	3372
4588	4489	2804	2844	1849	1854
2408	2424	4197	4127	5100	4972
7317	7290	1498	1474	4855	4668
894	900	2333	2322	786	738
3436	3528	18208	18271	2637	2661
628	<800	2298	2322	1590	1498
4548	4489	655	<800	1551	1498
10022	9975	4722	4639	2404	2419
1294	1270	3350	3349	2367	2419
3368	3404	1106	1120	7404	7763
2605	2636	6	<800	2550	2661
11049	11044	1252	1252	5821	5919
2901	3001	1865	1874	6442	6510
3810	3763	3820	3748	7628	7763
3224	3404	3738	3748	3409	3372
265	<800	1260	1252	5101	5174
1506	1489	2177	2158	4674	4668
2794	3001	11327	11212	1871	1854
10596	10633	8132	8134	3928	3943
5955	5803	8234	8134	3151	3177
2975	3001	9110	9098	10946	10956

7368	7290	7551	7596	1507	1498
740	<800	1891	1874	5891	5919
6924	6877	3698	3748	8949	9128
488	<800	218	<800	2678	2827
1689	1615	11309	11212	51	<800
634	<800	770	<800	6768	6510
376	<800	8011	8134	3884	3943
12088	12061	581	<800	2500	2661
		1780	1874	4918	4800
		641	<800	719	<800
		5216	5119	2412	2419
		2255	2322		
		1179	1120		
		734	<800		

```
FEATURES             Location/Qualifiers
     source          1. .165311
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-415L24"
                     /clone_lib="RPCI human BAC library 11"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 165311;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60695 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 60644


RESULT 139
AC135854/c
LOCUS       AC135854               165972 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 15 clone RP11-1404H5 map 15, 3 unordered
            pieces.
ACCESSION   AC135854
VERSION     AC135854.4  GI:27552578
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 165972)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-1404H5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 165972)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 165972)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 9, 2003 this sequence version replaced gi:27531844.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L28447
                Center clone name: 1404_H_5
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    22900: contig of 22900 bp in length
            *    22901    23000: gap of 100 bp
            *    23001   138676: contig of 115676 bp in length
            *   138677   138776: gap of 100 bp
            *   138777   165972: contig of 27196 bp in length.
FEATURES             Location/Qualifiers
     source          1. .165972
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-1404H5"
                     /clone_lib="RPCI-11 Male BAC segment 5"
     gap             22901. .23000
                     /estimated_length=100
     gap             138677. .138776
```

```
                     /estimated_length=100
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 165972;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76245 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 76194


RESULT 140
AC026830
LOCUS       AC026830              166484 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 15 clone RP11-215H14 map 15, WORKING DRAFT
            SEQUENCE, 32 unordered pieces.
ACCESSION   AC026830
VERSION     AC026830.2  GI:7596890
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 166484)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-215H14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 166484)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 166484)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 19, 2000 this sequence version replaced gi:7321633.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7706
                Center clone name: 215_H_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 150453 bases at least Q40
                Consensus quality: 158028 bases at least Q30
                Consensus quality: 161192 bases at least Q20
                Insert size: 176000; agarose-fp
                Insert size: 163384; sum-of-contigs
                Quality coverage: 3.6 in Q20 bases; agarose-fp
                Quality coverage: 3.9 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 32 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1024: contig of 1024 bp in length
            *     1025     1124: gap of 100 bp
            *     1125     2485: contig of 1361 bp in length
            *     2486     2585: gap of 100 bp
            *     2586     3626: contig of 1041 bp in length
            *     3627     3726: gap of 100 bp
            *     3727     4884: contig of 1158 bp in length
            *     4885     4984: gap of 100 bp
            *     4985     6513: contig of 1529 bp in length
            *     6514     6613: gap of 100 bp
            *     6614     8513: contig of 1900 bp in length
            *     8514     8613: gap of 100 bp
            *     8614    10616: contig of 2003 bp in length
            *    10617    10716: gap of 100 bp
            *    10717    12445: contig of 1729 bp in length
            *    12446    12545: gap of 100 bp
            *    12546    16043: contig of 3498 bp in length
            *    16044    16143: gap of 100 bp
            *    16144    19800: contig of 3657 bp in length
            *    19801    19900: gap of 100 bp
            *    19901    23755: contig of 3855 bp in length
            *    23756    23855: gap of 100 bp
            *    23856    25033: contig of 1178 bp in length
            *    25034    25133: gap of 100 bp
            *    25134    28296: contig of 3163 bp in length
            *    28297    28396: gap of 100 bp
            *    28397    31054: contig of 2658 bp in length
            *    31055    31154: gap of 100 bp
            *    31155    35420: contig of 4266 bp in length
            *    35421    35520: gap of 100 bp
            *    35521    39220: contig of 3700 bp in length
            *    39221    39320: gap of 100 bp
            *    39321    42676: contig of 3356 bp in length
            *    42677    42776: gap of 100 bp
            *    42777    46418: contig of 3642 bp in length
```

```
          *    46419    46518: gap of 100 bp
          *    46519    51021: contig of 4503 bp in length
          *    51022    51121: gap of 100 bp
          *    51122    56073: contig of 4952 bp in length
          *    56074    56173: gap of 100 bp
          *    56174    61048: contig of 4875 bp in length
          *    61049    61148: gap of 100 bp
          *    61149    66205: contig of 5057 bp in length
          *    66206    66305: gap of 100 bp
          *    66306    72086: contig of 5781 bp in length
          *    72087    72186: gap of 100 bp
          *    72187    77190: contig of 5004 bp in length
          *    77191    77290: gap of 100 bp
          *    77291    87414: contig of 10124 bp in length
          *    87415    87514: gap of 100 bp
          *    87515    94355: contig of 6841 bp in length
          *    94356    94455: gap of 100 bp
          *    94456   103495: contig of 9040 bp in length
          *   103496   103595: gap of 100 bp
          *   103596   112676: contig of 9081 bp in length
          *   112677   112776: gap of 100 bp
          *   112777   122142: contig of 9366 bp in length
          *   122143   122242: gap of 100 bp
          *   122243   131715: contig of 9473 bp in length
          *   131716   131815: gap of 100 bp
          *   131816   142646: contig of 10831 bp in length
          *   142647   142746: gap of 100 bp
          *   142747   166484: contig of 23738 bp in length.
FEATURES             Location/Qualifiers
     source          1. .166484
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-215H14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1024
                     /note="assembly_fragment"
     gap             1025. .1124
                     /estimated_length=100
     misc_feature    1125. .2485
                     /note="assembly_fragment"
     gap             2486. .2585
                     /estimated_length=100
     misc_feature    2586. .3626
                     /note="assembly_fragment"
     gap             3627. .3726
                     /estimated_length=100
     misc_feature    3727. .4884
                     /note="assembly_fragment"
     gap             4885. .4984
                     /estimated_length=100
     misc_feature    4985. .6513
                     /note="assembly_fragment"
     gap             6514. .6613
                     /estimated_length=100
     misc_feature    6614. .8513
                     /note="assembly_fragment"
     gap             8514. .8613
                     /estimated_length=100
     misc_feature    8614. .10616
                     /note="assembly_fragment"
     gap             10617. .10716
                     /estimated_length=100
     misc_feature    10717. .12445
                     /note="assembly_fragment"
     gap             12446. .12545

  Query Match             1.7%;  Score 52;  DB 14;  Length 166484;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67810 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 67861


RESULT 141
AL844892
LOCUS       AL844892             167780 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-396M20 on chromosome 10 Contains
            the 5' end of the GRID1 gene for glutamate receptor (ionotropic)
            delta 1, the 3' end of the gene for  friend of EBNA2 (FOE)
            (KIAA0261) and five CpG islands, complete sequence.
ACCESSION   AL844892
VERSION     AL844892.5  GI:23395595
KEYWORDS    HTG; FOE; GRID1; KIAA0261.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167780)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 1, 2002 this sequence version replaced gi:22798583.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-396M20 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .167780
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-396M20"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-396M20"
     misc_feature    48351. .48357
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     misc_feature    57532
                     /note="Clone_right_end: RP11-539A10"
     gene            join(complement(58796. .61348),
                     complement(AL683834.9:36029. .103858),
                     complement(AC073162.9:114306. .176049),
                     complement(AL596135.6:7440. .14051),
                     complement(AC022028.7:40751. .86154))
```

```
                     /gene="GRID1"
                     /locus_tag="RP11-93H12.1-001"
     mRNA            join(complement(61169. .61348),complement(58796. .58951),
                     complement(AL683834.9:103574. .103858),
                     complement(AL683834.9:36029. .36234),
                     complement(AC073162.9:175996. .176049),
                     complement(AC073162.9:128820. .128990),
                     complement(AC073162.9:115839. .116000),
                     complement(AC073162.9:114306. .114425),
                     complement(AL596135.6:13950. .14051),
                     complement(AL596135.6:12292. .12489),
                     complement(AL596135.6:8789. .9113),
                     complement(AL596135.6:7440. .7578),
                     complement(AC022028.7:85959. .86154),
                     complement(AC022028.7:52164. .52404),
                     complement(AC022028.7:40751. .41458))
                     /gene="GRID1"
                     /locus_tag="RP11-93H12.1-001"
                     /product="glutamate receptor, ionotropic, delta 1"
                     /note="match: cDNAs: BC039263.1"
     CDS             join(complement(61169. .61247),complement(58796. .58951),
                     complement(AL683834.9:103574. .103858),
                     complement(AL683834.9:36029. .36234),
                     complement(AC073162.9:175996. .176049),
                     complement(AC073162.9:128820. .128990),
                     complement(AC073162.9:115839. .116000),
                     complement(AC073162.9:114306. .114425),
                     complement(AL596135.6:13950. .14051),
                     complement(AL596135.6:12292. .12489),
                     complement(AL596135.6:8789. .9113),
                     complement(AL596135.6:7440. .7578),
                     complement(AC022028.7:85959. .86154),
                     complement(AC022028.7:52330. .52404))
                     /gene="GRID1"
                     /locus_tag="RP11-93H12.1-001"
                     /standard_name="OTTHUMP00000020002"
                     /note="match: proteins: Q61627 Q62640"
                     /codon_start=1
                     /product="glutamate receptor, ionotropic, delta 1"
                     /protein_id="CAH73378.1"
                     /db_xref="GI:55666097"
                     /db_xref="GOA:Q5VSK3"
                     /db_xref="InterPro:IPR001311"
                     /db_xref="InterPro:IPR001320"
                     /db_xref="InterPro:IPR001508"
                     /db_xref="InterPro:IPR001828"
                     /db_xref="UniProt/TrEMBL:Q5VSK3"
                     /translation="MEALTLWLLPWICQCVSVRADSIIHIGAIFEENAAKDDRVFQLA
                     VSDLSLNDDILQSEKITYSIKVIEANNPFQAVQEACDLMTQGILALVTSTGCASANAL
                     QSLTDAMHIPHLFVQRNPGGSPRTACHLNPSPDGEAYTLASRPPVRLNDVMLRLVTEL
                     RWQKFVMFYDSEYDIRGLQSFLDQASRLGLDVSLQKVDKNISHVFTSLFTTMKTEELN
                     RYRDTLRRAILLLSPQGAHSFINEAVETNLASKDSHWVFVNEEISDPEILDLVHSALG
                     RMTVVRQIFPSAKDNQKCTRNNHRISSLLCDPQEGYLQMLQISNLYLYDSVLMLANAF
                     HRKLEDRKWHSMASLNCIRKSTKPWNGGRSMLDTIKKGHITGLTGVMEFREDSSNPYV
                     QFEILGTTYSETFGKDMRKLATWDSEKGLNGSLQERPMGSRLQGLTLKVVTVLEEPFV
                     MVAENILGQPKRYKGFSIDVLDALAKALGFKYEIYQAPDGRYGHQLHNTSWNGMIGEL
                     ISKRADLAISAITITPERESVVDFSKRYMDYSVGILIKKPEEKISIFSLFAPFDFAVW
                     ACIAAAIPVVGVLIFVLNRIQAVRAQSAAQPRPSASATLHSAIWIVYGAFVQQGGESS
                     VNSMAMRIVMGSWWLFTLIVCSSYTANLAAFLTVSRMDNPIRTFQDLSKQVEMSYGTV
                     RDSAVYEYFRAKGTNPLEQDSTFAELWRTISKNGGADNCVSSPSEGIRKDPGAAGHRG
                     PGCAEAEVVAAHGPL"
     misc_feature    61673. .61783
                     /note="Single clone region. Sequence generated from a
                     transposon library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     gene            join(complement(AL731569.18:50750. .50890),
                     complement(AL731569.18:50440. .50616),
                     complement(AL731569.18:46646. .47166),
                     complement(AL731569.18:28793. .29818),
                     complement(AL731569.18:26276. .26394),
                     complement(AL731569.18:2953. .3048),
                     complement(167467. .167619),complement(167063. .167199),
                     complement(165847. .165958),complement(162134. .162361),
                     complement(156031. .156142),complement(155254. .155351),
                     complement(153844. .153920),complement(148511. .148686),
                     complement(148075. .148190),complement(146822. .146986),
                     complement(141097. .141304),complement(138130. .138218),
                     complement(132775. .132870),complement(130111. .132463))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-001"
     mRNA            join(complement(AL731569.18:50750. .50890),
                     complement(AL731569.18:50440. .50616),
                     complement(AL731569.18:46646. .47166),
                     complement(AL731569.18:28793. .29818),
                     complement(AL731569.18:26276. .26394),
                     complement(AL731569.18:2953. .3048),
                     complement(167467. .167619),complement(167063. .167199),
                     complement(165847. .165958),complement(162134. .162361),
                     complement(156031. .156142),complement(155254. .155351),
                     complement(153844. .153920),complement(148511. .148686),
                     complement(148075. .148190),complement(146822. .146986),
                     complement(141097. .141304),complement(138130. .138218),
                     complement(132775. .132870),complement(130111. .132463))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-001"
                     /product="KIAA0261"
                     /note="match: ESTs: AA579854.1 AW577904.1 BI438658.1
                     BI459087.1 BM504319.1 BM683407.1 BQ007553.1 BQ420406.1
                     BU620920.1
                     match: cDNAs: D87450.1"
     gene            complement(join(130111. .132463,132775. .132870,
                     138130. .138887))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-002"
     mRNA            complement(join(130111. .132463,132775. .132870,
                     138130. .138887))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-002"
                     /product="KIAA0261"
                     /note="match: cDNAs: BC036507.1"
     polyA_site      complement(130111)
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-002"
     polyA_signal    complement(130129. .130134)
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-002"
     gene            complement(join(131865. .132463,132775. .132870,
                     138130. .138218,141097. .141304,148075. .148190,
                     148511. .148686,153844. .153920,155254. .155351,
                     156031. .156142,162134. .162363))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-004"
     mRNA            complement(join(131865. .132463,132775. .132870,
                     138130. .138218,141097. .141304,148075. .148190,
                     148511. .148686,153844. .153920,155254. .155351,
                     156031. .156142,162134. .162363))
                     /gene="KIAA0261"
                     /locus_tag="RP11-396M20.1-004"
                     /product="KIAA0261"
                     /note="match: cDNAs: BC017393.1"
     CDS             join(complement(AL731569.18:50750. .50805),
                     complement(AL731569.18:50440. .50616),
                     complement(AL731569.18:46646. .47166),
                     complement(AL731569.18:28793. .29818),

  Query Match             1.7%;  Score 52;  DB 8;  Length 167780;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93241 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 93292


RESULT 142
AC109635
```

```
LOCUS       AC109635               167920 bp    DNA     linear   PRI 10-SEP-2002
DEFINITION  Homo sapiens chromosome 11, clone RP11-347H15, complete sequence.
ACCESSION   AC109635
VERSION     AC109635.6  GI:22773373
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167920)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-347H15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 167920)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 167920)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 167920)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 10, 2002 this sequence version replaced gi:22475361.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L24041
                Center clone name: 347_H_15
            --------------.
FEATURES             Location/Qualifiers
     source          1. .167920
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-347H15"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(903. .2453)
                     /rpt_family="pTR5"
     repeat_region   2562. .2726
                     /rpt_family="MIR"
     repeat_region   complement(3163. .3289)
                     /rpt_family="AluJb"
     repeat_region   3377. .3426
                     /rpt_family="(CATA)n"
     repeat_region   3701. .5177
                     /rpt_family="THE1-int"
     repeat_region   5409. .5446
                     /rpt_family="(CA)n"
     repeat_region   6146. .6179
                     /rpt_family="(TTTG)n"
     repeat_region   complement(6551. .6861)
                     /rpt_family="AluSx"
     repeat_region   6868. .6965
                     /rpt_family="MIR"
     repeat_region   9236. .9316
                     /rpt_family="L2"
     repeat_region   9478. .9544
                     /rpt_family="(TCTA)n"
     repeat_region   complement(10278. .10409)
                     /rpt_family="FLAM_C"
     repeat_region   10439. .10729
                     /rpt_family="AluY"
     repeat_region   complement(10738. .10947)
                     /rpt_family="L1ME"
     repeat_region   complement(10968. .11095)
                     /rpt_family="L1ME4A"
     repeat_region   complement(11142. .11423)
                     /rpt_family="L1ME4A"
     repeat_region   11462. .11535
                     /rpt_family="(TATATG)n"
     repeat_region   11813. .11846
                     /rpt_family="AT_rich"
     repeat_region   complement(13232. .13345)
                     /rpt_family="FRAM/FAM"
     repeat_region   13679. .13724
                     /rpt_family="(TATATG)n"
```

```
     repeat_region   13725. .13893
                     /rpt_family="(TA)n"
     repeat_region   13898. .14082
                     /rpt_family="(TAA)n"
     repeat_region   14083. .14248
                     /rpt_family="(TA)n"
     repeat_region   14272. .14947
                     /rpt_family="(TA)n"
     repeat_region   14990. .15155
                     /rpt_family="(TA)n"
     repeat_region   15202. .15371
                     /rpt_family="(TA)n"
     repeat_region   15378. .15398
                     /rpt_family="AT_rich"
     repeat_region   15411. .15730
                     /rpt_family="(TA)n"
     repeat_region   15735. .15761
                     /rpt_family="AT_rich"
     repeat_region   15766. .15926
                     /rpt_family="(TA)n"
     repeat_region   15942. .16119
                     /rpt_family="(TA)n"
     unsure          16041. .16047
                     /note="<30 qual SNGL region"
     repeat_region   complement(16537. .16909)
                     /rpt_family="L2"
     repeat_region   16930. .16965
                     /rpt_family="(TG)n"
     repeat_region   17126. .17420
                     /rpt_family="AluSx"
     repeat_region   complement(17423. .17566)
                     /rpt_family="L1PA2"
     repeat_region   17567. .18711
                     /rpt_family="L1PA2"
     repeat_region   complement(19966. .20063)
                     /rpt_family="MIR"
     repeat_region   complement(20925. .20973)
                     /rpt_family="L1MC4"
     repeat_region   20977. .22553
                     /rpt_family="pTR5"
     repeat_region   complement(22640. .23478)
                     /rpt_family="L1MC4"
     repeat_region   23560. .23902
                     /rpt_family="L2"
     repeat_region   complement(23943. .24673)
                     /rpt_family="LTR8"
     repeat_region   24699. .24746
                     /rpt_family="(CGTG)n"
     repeat_region   25464. .25498
                     /rpt_family="MADE1"
     repeat_region   26408. .26725
                     /rpt_family="L1PA7"
     repeat_region   26726. .26765
                     /rpt_family="AT_rich"

  Query Match             1.7%;  Score 52;  DB 8;  Length 167920;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57209 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 57260


RESULT 143
AC007371
LOCUS       AC007371               168734 bp    DNA     linear   PRI 16-SEP-1999
DEFINITION  Homo sapiens, clone 24_A_9, complete sequence.
ACCESSION   AC007371
VERSION     AC007371.16  GI:5542041
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 168734)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone 24_A_9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 168734)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 168734)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K.,
            Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P.,
            FitzHugh,W., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Klein,J., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 21, 1999 this sequence version replaced gi:5468556.
            All repeats were identified using RepeatMasker: Smit, A.F.A. &
            Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES             Location/Qualifiers
     source          1. .168734
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="24_A_9"
                     /clone_lib="Alan Buckler -- per comm"
     repeat_region   complement(1444. .1552)
                     /rpt_family="MIR"
     repeat_region   complement(1581. .1792)
                     /rpt_family="MER20"
     repeat_region   complement(3654. .3953)
                     /rpt_family="AluSg"
     repeat_region   4172. .4695
                     /rpt_family="L1MC4"
     repeat_region   4700. .4730
                     /rpt_family="(TTTG)n"
     repeat_region   complement(4731. .5015)
                     /rpt_family="AluSq"
     repeat_region   5016. .5038
                     /rpt_family="L1MC4"
     repeat_region   complement(5039. .5331)
                     /rpt_family="AluY"
     repeat_region   5332. .5417
                     /rpt_family="L1MC4"
     repeat_region   5490. .5649
```

```
                     /rpt_family="MIR"
     repeat_region   6091. .6249
                     /rpt_family="MIR"
     repeat_region   6615. .6645
                     /rpt_family="AT_rich"
     repeat_region   6920. .7228
                     /rpt_family="AluSg"
     repeat_region   complement(7875. .8062)
                     /rpt_family="L2"
     repeat_region   8069. .8161
                     /rpt_family="MIR"
     repeat_region   complement(10637. .10816)
                     /rpt_family="L1MC4"
     repeat_region   complement(11031. .11134)
                     /rpt_family="MER112"
     repeat_region   11375. .11507
                     /rpt_family="FLAM_C"
     repeat_region   11901. .11932
                     /rpt_family="AT_rich"
     repeat_region   12275. .12296
                     /rpt_family="AT_rich"
     repeat_region   12424. .12632
                     /rpt_family="MIR"
     repeat_region   12877. .13606
                     /rpt_family="L2"
     repeat_region   14281. .14308
                     /rpt_family="AT_rich"
     repeat_region   14414. .15069
                     /rpt_family="L2"
     repeat_region   15459. .15589
                     /rpt_family="FLAM_C"
     repeat_region   15650. .16082
                     /rpt_family="L2"
     repeat_region   16280. .16311
                     /rpt_family="(TTTG)n"
     repeat_region   complement(16862. .17059)
                     /rpt_family="MIR"
     repeat_region   17113. .17366
                     /rpt_family="MIR"
     repeat_region   17405. .17439
                     /rpt_family="(TG)n"
     repeat_region   17591. .17645
                     /rpt_family="(CA)n"
     repeat_region   complement(18182. .18482)
                     /rpt_family="AluSx"
     repeat_region   complement(19280. .19588)
                     /rpt_family="AluJb"
     repeat_region   19617. .19903
                     /rpt_family="AluJo"
     repeat_region   19904. .19934
                     /rpt_family="(CAA)n"
     repeat_region   20626. .20927
                     /rpt_family="AluJb"
     repeat_region   21346. .21380
                     /rpt_family="AT_rich"
     repeat_region   22056. .22353
                     /rpt_family="AluSg"
     repeat_region   22650. .22705
                     /rpt_family="L2"
     repeat_region   complement(22705. .22823)
                     /rpt_family="MIR"
     repeat_region   24066. .24223
                     /rpt_family="MLT1J2"
     repeat_region   24294. .24435
                     /rpt_family="MLT1J1"
     repeat_region   complement(24436. .24606)
                     /rpt_family="AluSg/x"
     repeat_region   24646. .24717
                     /rpt_family="L2"
     repeat_region   24754. .24994
                     /rpt_family="MLT1B"
     repeat_region   25223. .25277
                     /rpt_family="MIR"
     repeat_region   complement(25440. .25546)
                     /rpt_family="MIR"
     repeat_region   25616. .25690
                     /rpt_family="MER103"
     repeat_region   25885. .26015
                     /rpt_family="AluSq/x"
     repeat_region   complement(26034. .26094)
                     /rpt_family="MIR"
     repeat_region   26129. .26303
                     /rpt_family="MIR"
     repeat_region   26704. .26960
                     /rpt_family="AluSg1"
     repeat_region   complement(27595. .27889)
                     /rpt_family="AluSx"
     repeat_region   28051. .28369
                     /rpt_family="AluSx"
     repeat_region   28869. .28914
                     /rpt_family="AT_rich"
     repeat_region   29087. .29157
                     /rpt_family="MER47"
     repeat_region   30863. .31007
                     /rpt_family="MIR"
     repeat_region   complement(31619. .31897)
                     /rpt_family="AluJb"
     repeat_region   complement(31920. .32222)
                     /rpt_family="AluSg1"
     repeat_region   complement(32390. .32635)
                     /rpt_family="L2"
     repeat_region   complement(33248. .33322)
                     /rpt_family="L2"
     repeat_region   complement(33330. .33542)
                     /rpt_family="MIR"
     repeat_region   33762. .33782
                     /rpt_family="AT_rich"
     repeat_region   35250. .35296
                     /rpt_family="AT_rich"
     repeat_region   35406. .35454
                     /rpt_family="(TG)n"
     repeat_region   complement(35466. .35758)
                     /rpt_family="AluSx"
     repeat_region   38644. .38729
                     /rpt_family="L2"
     repeat_region   38842. .38868
                     /rpt_family="(TTTC)n"
     repeat_region   38873. .38937
                     /rpt_family="L2"
     repeat_region   39359. .39411

  Query Match             1.7%;  Score 52;  DB 8;  Length 168734;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28288 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 28339


RESULT 144
AC025787
LOCUS       AC025787              168793 bp    DNA     linear   PRI 20-FEB-2002
DEFINITION  Homo sapiens chromosome 11, clone RP11-227P3, complete sequence.
ACCESSION   AC025787
VERSION     AC025787.9  GI:18699972
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 168793)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-227P3
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 168793)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 168793)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 168793)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 18, 2002 this sequence version replaced gi:18377313.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6819
                Center clone name: 227_P_3
            --------------.
FEATURES             Location/Qualifiers
     source          1. .168793
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-227P3"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(1. .97)
                     /rpt_family="L2"
     repeat_region   100. .393
                     /rpt_family="AluJb"
     repeat_region   complement(648. .1118)
                     /rpt_family="MER1A"
     repeat_region   complement(1420. .1573)
                     /rpt_family="LTR41"
     repeat_region   1574. .1717
                     /rpt_family="AluJo"
     repeat_region   1718. .1757
                     /rpt_family="(TAAAA)n"
     repeat_region   1758. .1925
                     /rpt_family="AluJo"
     repeat_region   2227. .2530
                     /rpt_family="AluSq"
     repeat_region   complement(2540. .2677)
                     /rpt_family="MIR"
     repeat_region   3942. .3968
                     /rpt_family="AT_rich"
     repeat_region   4486. .4792
                     /rpt_family="AluSq"
     repeat_region   complement(5142. .5451)
                     /rpt_family="AluSx"
     repeat_region   complement(5848. .6129)
                     /rpt_family="AluY"
     repeat_region   complement(6442. .6567)
                     /rpt_family="MER117"
     repeat_region   7601. .7893
                     /rpt_family="AluSx"
     repeat_region   8001. .8296
                     /rpt_family="AluSg"
     repeat_region   8385. .8663
                     /rpt_family="AluY"
     unsure          complement(8640. .8645)
                     /note="<30 qual SNGL region"
     repeat_region   complement(8840. .9137)
                     /rpt_family="AluSx"
     repeat_region   complement(9183. .9481)
                     /rpt_family="AluSx"
     repeat_region   9630. .9754
                     /rpt_family="AT_rich"
     repeat_region   9861. .9889
                     /rpt_family="AluJo"
     repeat_region   complement(9894. .10195)
                     /rpt_family="AluSx"
     repeat_region   10196. .10467
```

```
                     /rpt_family="AluJo"
     repeat_region   10876. .11186
                     /rpt_family="L2"
     unsure          11226. .11229
                     /note="probably TTTT; possibly TTT"
     repeat_region   complement(11286. .11583)
                     /rpt_family="AluSg"
     repeat_region   11621. .11699
                     /rpt_family="L2"
     repeat_region   complement(11700. .11920)
                     /rpt_family="MER58A"
     repeat_region   11921. .12456
                     /rpt_family="L2"
     repeat_region   12523. .12649
                     /rpt_family="AluJb"
     repeat_region   12652. .12866
                     /rpt_family="AluJo"
     repeat_region   12870. .13185
                     /rpt_family="AluSx"
     repeat_region   13191. .13308
                     /rpt_family="L2"
     repeat_region   complement(13535. .13839)
                     /rpt_family="AluSg"
     repeat_region   15032. .15436
                     /rpt_family="L1MC5"
     repeat_region   15459. .15550
                     /rpt_family="L1MC5"
     repeat_region   15597. .15869
                     /rpt_family="AluSq"
     repeat_region   complement(15870. .20120)
                     /rpt_family="ALR6/alpha"
     repeat_region   20121. .20182
                     /rpt_family="(TA)n"
     repeat_region   complement(20183. .21439)
                     /rpt_family="ALR6/alpha"
     repeat_region   21996. .22132
                     /rpt_family="MIR3"
     repeat_region   complement(22855. .23061)
                     /rpt_family="L2"
     unsure          complement(22904. .22918)

  Query Match             1.7%;  Score 52;  DB 8;  Length 168793;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121606 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 121657


RESULT 145
AC009790/c
LOCUS       AC009790              169059 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 8 clone RP11-128N8 map 8, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC009790
VERSION     AC009790.4  GI:7528038
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 169059)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-128N8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169059)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 169059)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 8, 2000 this sequence version replaced gi:6630600.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2031
                Center clone name: 128_N_8
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 139988 bases at least Q40
                Consensus quality: 156241 bases at least Q30
                Consensus quality: 162235 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 166659; sum-of-contigs
                Quality coverage: 3.9 in Q20 bases; agarose-fp
                Quality coverage: 4.0 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1539: contig of 1539 bp in length
            *     1540     1639: gap of 100 bp
            *     1640     2821: contig of 1182 bp in length
```

```
            *     2822     2921: gap of 100 bp
            *     2922     4692: contig of 1771 bp in length
            *     4693     4792: gap of 100 bp
            *     4793     6136: contig of 1344 bp in length
            *     6137     6236: gap of 100 bp
            *     6237     8728: contig of 2492 bp in length
            *     8729     8828: gap of 100 bp
            *     8829    11075: contig of 2247 bp in length
            *    11076    11175: gap of 100 bp
            *    11176    13885: contig of 2710 bp in length
            *    13886    13985: gap of 100 bp
            *    13986    17062: contig of 3077 bp in length
            *    17063    17162: gap of 100 bp
            *    17163    21818: contig of 4656 bp in length
            *    21819    21918: gap of 100 bp
            *    21919    25663: contig of 3745 bp in length
            *    25664    25763: gap of 100 bp
            *    25764    30441: contig of 4678 bp in length
            *    30442    30541: gap of 100 bp
            *    30542    33886: contig of 3345 bp in length
            *    33887    33986: gap of 100 bp
            *    33987    38474: contig of 4488 bp in length
            *    38475    38574: gap of 100 bp
            *    38575    43895: contig of 5321 bp in length
            *    43896    43995: gap of 100 bp
            *    43996    50044: contig of 6049 bp in length
            *    50045    50144: gap of 100 bp
            *    50145    56692: contig of 6548 bp in length
            *    56693    56792: gap of 100 bp
            *    56793    61378: contig of 4586 bp in length
            *    61379    61478: gap of 100 bp
            *    61479    69459: contig of 7981 bp in length
            *    69460    69559: gap of 100 bp
            *    69560    78975: contig of 9416 bp in length
            *    78976    79075: gap of 100 bp
            *    79076    88285: contig of 9210 bp in length
            *    88286    88385: gap of 100 bp
            *    88386    98036: contig of 9651 bp in length
            *    98037    98136: gap of 100 bp
            *    98137   109264: contig of 11128 bp in length
            *   109265   109364: gap of 100 bp
            *   109365   123064: contig of 13700 bp in length
            *   123065   123164: gap of 100 bp
            *   123165   143558: contig of 20394 bp in length
            *   143559   143658: gap of 100 bp
            *   143659   169059: contig of 25401 bp in length.
FEATURES             Location/Qualifiers
     source          1. .169059
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-128N8"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1539
                     /note="assembly_fragment"
     gap             1540. .1639
                     /estimated_length=100
     misc_feature    1640. .2821
                     /note="assembly_fragment"
     gap             2822. .2921
                     /estimated_length=100
     misc_feature    2922. .4692
                     /note="assembly_fragment"
     gap             4693. .4792
                     /estimated_length=100
     misc_feature    4793. .6136
                     /note="assembly_fragment"
     gap             6137. .6236
                     /estimated_length=100
     misc_feature    6237. .8728
                     /note="assembly_fragment"
     gap             8729. .8828
                     /estimated_length=100
     misc_feature    8829. .11075
                     /note="assembly_fragment"
     gap             11076. .11175
                     /estimated_length=100
     misc_feature    11176. .13885
                     /note="assembly_fragment"
     gap             13886. .13985
                     /estimated_length=100
     misc_feature    13986. .17062
                     /note="assembly_fragment"
     gap             17063. .17162
                     /estimated_length=100
     misc_feature    17163. .21818
                     /note="assembly_fragment"
     gap             21819. .21918
                     /estimated_length=100
     misc_feature    21919. .25663
                     /note="assembly_fragment"
     gap             25664. .25763
                     /estimated_length=100
     misc_feature    25764. .30441
                     /note="assembly_fragment"
     gap             30442. .30541
                     /estimated_length=100
     misc_feature    30542. .33886
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     gap             33887. .33986
                     /estimated_length=100
     misc_feature    33987. .38474

  Query Match             1.7%;  Score 52;  DB 14;  Length 169059;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     158812 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 158761


RESULT 146
AC021165/c
LOCUS       AC021165               169893 bp    DNA     linear   HTG 02-SEP-2000
DEFINITION  Homo sapiens chromosome 19 clone RP11-565J3, WORKING DRAFT
            SEQUENCE, 17 unordered pieces.
ACCESSION   AC021165
VERSION     AC021165.4  GI:9965022
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 169893)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169893)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Sep 2, 2000 this sequence version replaced gi:7637373.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
```

```
           Center project name: H_NH0565J03
           ------------- Summary Statistics  --------------
           Sequencing vector: M13; 76%
           Sequencing vector: plasmid; 24%
           Chemistry: Dye-primer ET; 76% of reads
           Chemistry: Dye-terminator Big Dye; 24% of reads
           Assembly program: Phrap; version 0.990319
           Consensus quality: 152885 bases at least Q40
           Consensus quality: 159085 bases at least Q30
           Consensus quality: 162352 bases at least Q20
           Insert size: 185000; agarose-fp
           Insert size: 168293; sum-of-contigs
           Quality coverage: 3.07 in Q20 bases; agarose-fp
           Quality coverage: 3.43 in Q20 bases; sum-of-contigs
           ------------------------------------------------.
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 17 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *        1     1397: contig of 1397 bp in length
           *     1398     1497: gap of unknown length
           *     1498     3037: contig of 1540 bp in length
           *     3038     3137: gap of unknown length
           *     3138     6455: contig of 3318 bp in length
           *     6456     6555: gap of unknown length
           *     6556     9394: contig of 2839 bp in length
           *     9395     9494: gap of unknown length
           *     9495    12928: contig of 3434 bp in length
           *    12929    13028: gap of unknown length
           *    13029    17141: contig of 4113 bp in length
           *    17142    17241: gap of unknown length
           *    17242    23019: contig of 5778 bp in length
           *    23020    23119: gap of unknown length
           *    23120    27296: contig of 4177 bp in length
           *    27297    27396: gap of unknown length
           *    27397    53901: contig of 26505 bp in length
           *    53902    54001: gap of unknown length
           *    54002    61627: contig of 7626 bp in length
           *    61628    61727: gap of unknown length
           *    61728    72148: contig of 10421 bp in length
           *    72149    72248: gap of unknown length
           *    72249    80663: contig of 8415 bp in length
           *    80664    80763: gap of unknown length
           *    80764    89249: contig of 8486 bp in length
           *    89250    89349: gap of unknown length
           *    89350   102831: contig of 13482 bp in length
           *   102832   102931: gap of unknown length
           *   102932   115578: contig of 12647 bp in length
           *   115579   115678: gap of unknown length
           *   115679   131546: contig of 15868 bp in length
           *   131547   131646: gap of unknown length
           *   131647   169893: contig of 38247 bp in length.
FEATURES             Location/Qualifiers
.    source          1. .169893
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="RP11-565J3"
     misc_feature    1. .1397
                     /note="assembly_name:Contig10"
     gap             1398. .1497
                     /estimated_length=unknown
     misc_feature    1498. .3037
                     /note="assembly_name:Contig11"
     gap             3038. .3137
                     /estimated_length=unknown
     misc_feature    3138. .6455
                     /note="assembly_name:Contig12"
     gap             6456. .6555
                     /estimated_length=unknown
     misc_feature    6556. .9394
                     /note="assembly_name:Contig13"
     gap             9395. .9494
                     /estimated_length=unknown
     misc_feature    9495. .12928
                     /note="assembly_name:Contig14"
     gap             12929. .13028
                     /estimated_length=unknown
     misc_feature    13029. .17141
                     /note="assembly_name:Contig15"
     gap             17142. .17241
                     /estimated_length=unknown
     misc_feature    17242. .23019
                     /note="assembly_name:Contig16"
     gap             23020. .23119
                     /estimated_length=unknown
     misc_feature    23120. .27296
                     /note="assembly_name:Contig17
                     clone_end:SP6
                     vector_side:left"
     gap             27297. .27396
                     /estimated_length=unknown
     misc_feature    27397. .53901
                     /note="assembly_name:Contig25
                     clone_end:T7
                     vector_side:left"
     gap             53902. .54001
                     /estimated_length=unknown
     misc_feature    54002. .61627
                     /note="assembly_name:Contig18"
     gap             61628. .61727
                     /estimated_length=unknown
     misc_feature    61728. .72148
                     /note="assembly_name:Contig19"
     gap             72149. .72248
                     /estimated_length=unknown
     misc_feature    72249. .80663
                     /note="assembly_name:Contig20"
     gap             80664. .80763
                     /estimated_length=unknown
     misc_feature    80764. .89249
                     /note="assembly_name:Contig21"
     gap             89250. .89349
                     /estimated_length=unknown
     misc_feature    89350. .102831
                     /note="assembly_name:Contig22"
     gap             102832. .102931
                     /estimated_length=unknown
     misc_feature    102932. .115578
                     /note="assembly_name:Contig23"
     gap             115579. .115678
                     /estimated_length=unknown
     misc_feature    115679. .131546
                     /note="assembly_name:Contig24"
     gap             131547. .131646
                     /estimated_length=unknown
     misc_feature    131647. .169893
                     /note="assembly_name:Contig26"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 169893;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32826 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 32775


RESULT 147
AL135903/c
```

```
LOCUS       AL135903              170232 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-30P6 on chromosome 6 Contains a
            karyopherin alpha 5 (importin alpha 6) (KPNA5) pseudogene, a
            keratin 18 (KRT18) pseudogene, a tumor protein,
            translationally-controlled 1 (TPT1) pseudogene, a pseudogene
            similar to dUTP pyrophosphatase (DUT), the NT5E gene for
            5'-nucleotidase ecto (CD73) and a CpG island, complete sequence.
ACCESSION   AL135903
VERSION     AL135903.12  GI:12830388
KEYWORDS    HTG; CD73; CpG island; DUT; KPNA5; KRT18; NT5E; TPT1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 170232)
  AUTHORS   Johnson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 14, 2001 this sequence version replaced gi:11342762.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-30P6 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .170232
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-30P6"
                     /clone_lib="RPCI-11.1"
     misc_feature    104
                     /note="Clone_right_end: RP1-161C16"
     gene            join(520. .610,682. .1243,1488. .2119,4690. .4777,5077. .5293)
                     /locus_tag="RP11-30P6.1-001"
                     /pseudo
     CDS             join(520. .610,682. .1243,1488. .2119,4690. .4777,5077. .5293)
                     /locus_tag="RP11-30P6.1-001"
                     /note="match: proteins: AAL87378 IMA1_SCHPO IMA2_ARATH
                     IMA3_HUMAN IMA_LYCES K1M8_HUMAN K2C8_MOUSE O49601 O82783
                     O94374 Q90YH3 Q94F55 Q9D637 Q9PUB5 Q9U1H6"
                     /pseudo
                     /codon_start=1
     gene            2233. .3451
                     /locus_tag="RP11-30P6.2-001"
                     /pseudo
     CDS             2233. .3451
                     /locus_tag="RP11-30P6.2-001"
                     /note="match: proteins: AAH04253 AAH20982 K1CR_HUMAN
                     K1CR_MOUSE K1M8_HUMAN K2C8_MOUSE O57611 O77727 Q07427
                     Q90W74 Q90YH3 Q96GD2 Q9PUB5 Q9PWD8 Q9QWL7"
                     /pseudo
                     /codon_start=1
     gene            complement(join(101663. .104001,104457. .104630))
                     /locus_tag="RP11-30P6.6-001"
     mRNA            complement(join(101663. .104001,104457. .104630))
                     /locus_tag="RP11-30P6.6-001"
                     /note="match: cDNAs: AK024998"
     gene            complement(141101. .141569)
                     /locus_tag="RP11-30P6.4-001"
                     /pseudo
     CDS             complement(141101. .141569)
                     /locus_tag="RP11-30P6.4-001"
                     /note="match: proteins: AAB96413 AAD50334 AAL77806
                     CAB54626 DUT_ADEG1 DUT_ADEG8 DUT_FOWPV DUT_HUMAN DUT_ORFN2
                     DUT_RAT DUT_SCHPO DUT_VACCC DUT_VARV DUT_YEAST O01934
                     P87630 Q8V546 Q8VCG1 Q91BK5 Q96Q81 Q99HX3 Q9CQ43 Q9CU90
                     Q9JFF1 Q9JJ44 Q9Q8S9 Q9Q958 Q9STG6 Q9UI74"
                     /pseudo
                     /codon_start=1
     gene            complement(141897. .142410)
                     /locus_tag="RP11-30P6.3-001"
                     /pseudo
     CDS             complement(141897. .142410)
                     /locus_tag="RP11-30P6.3-001"
                     /note="match: proteins: AAH22436 AAL68965 FKG2_HUMAN
                     Q9NXW1 TCTP_BRABE TCTP_CHICK TCTP_DROME TCTP_HUMAN
                     TCTP_LABRO TCTP_MEDSA TCTP_MOUSE TCTP_PEA TCTP_RABIT
                     TCTP_SCHJA TCTP_SCHPO TCTP_YEAST"
                     /pseudo
                     /codon_start=1
     gene            join(164535. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15737,AL589666.5:29547. .29744,
                     AL589666.5:31647. .31801,AL589666.5:33806. .33911,
                     AL589666.5:34820. .34969,AL589666.5:36289. .36489,
                     AL589666.5:38153. .40090)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-001"
     gene            join(164535. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15911)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-002"
     mRNA            join(164535. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15911)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-001"
                     /product="5'-nucleotidase, ecto (CD73)"
                     /note="match: ESTs: AI202304 AI420969 AW006723 AW025554
                     AW271686 BE677446 BF476532
                     match: cDNAs: BC015940 BC015940.1"
     mRNA            join(164535. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15737,AL589666.5:29547. .29744,
                     AL589666.5:31647. .31801,AL589666.5:33806. .33911,
                     AL589666.5:34820. .34969,AL589666.5:36289. .36489,
                     AL589666.5:38153. .40090)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-001"
                     /product="5'-nucleotidase, ecto (CD73)"
                     /note="match: ESTs: AA375089 AL553737 AL702918 BB646149
                     BE547456 BE891902 BG294369 BI756281 BM006258 BM762339
                     BM847311 BM926263 BM937597
                     match: cDNAs: BT5NUC DO5NUCLT HS5NUASE MME5NA RN5RPHP
                     S64302"
     CDS             join(164584. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15781)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-001"
                     /standard_name="OTTHUMP00000016809"
                     /note="match: proteins: Q96B60"
                     /codon_start=1
                     /product="5'-nucleotidase, ecto (CD73)"
```

```
                     /protein_id="CAH72336.1"
                     /db_xref="GI:55661713"
                     /db_xref="GOA:Q96B60"
                     /db_xref="InterPro:IPR004843"
                     /db_xref="InterPro:IPR006146"
                     /db_xref="InterPro:IPR006179"
                     /db_xref="UniProt/TrEMBL:Q96B60"
                     /translation="MCPRAARAPATLLLALGAVLWPAAGAWELTILHTNDVHSRLEQT
                     SEDSSKCVNASRCMGGVARLFTKVQQIRRAEPNVLLLDAGDQYQGTIWFTVYKGAEVA
                     HFMNALRYDAMALGNHEFDNGVEGLIEPLLKEAKFPILSANIKAKGPLASQISGLYLP
                     YKVLPVGDEVVGIVGYTSKETPFLSNPGTNLVFEDEITALQPEVDKLKTLNVNKIIAL
                     GHSGFEMDKLIAQKVRGVDVVVGGHSNTFLYTGNCFKRIAWARMSR"
     CDS             join(164584. .164922,AL589666.5:11372. .11594,
                     AL589666.5:15549. .15737,AL589666.5:29547. .29744,
                     AL589666.5:31647. .31801,AL589666.5:33806. .33911,
                     AL589666.5:34820. .34969,AL589666.5:36289. .36489,
                     AL589666.5:38153. .38316)
                     /gene="NT5E"
                     /locus_tag="RP11-321N4.1-001"
                     /standard_name="OTTHUMP00000016808"
                     /note="match: proteins: 5NTD_BOOMI 5NTD_BOVIN 5NTD_DISOM
                     5NTD_HAEIN 5NTD_HUMAN 5NTD_LUTLO 5NTD_MOUSE 5NTD_RAT
                     AAK89637 AAL44580 APY_AEDAE CAD28125 CAD28126 O34313
                     Q05700 Q95P65 Q96B60 Q9U9I6 Q9V824 Q9VZ33"
                     /codon_start=1
                     /product="5'-nucleotidase, ecto (CD73)"
                     /protein_id="CAH72337.1"
                     /db_xref="GI:55661714"
                     /db_xref="InterPro:IPR004843"
                     /db_xref="InterPro:IPR006146"
                     /db_xref="InterPro:IPR006179"
                     /db_xref="InterPro:IPR008334"
                     /translation="MCPRAARAPATLLLALGAVLWPAAGAWELTILHTNDVHSRLEQT
                     SEDSSKCVNASRCMGGVARLFTKVQQIRRAEPNVLLLDAGDQYQGTIWFTVYKGAEVA
                     HFMNALRYDAMALGNHEFDNGVEGLIEPLLKEAKFPILSANIKAKGPLASQISGLYLP
                     YKVLPVGDEVVGIVGYTSKETPFLSNPGTNLVFEDEITALQPEVDKLKTLNVNKIIAL
                     GHSGFEMDKLIAQKVRGVDVVVGGHSNTFLYTGNPPSKEVPAGKYPFIVTSDDGRKVP
                     VVQAYAFGKYLGYLKIEFDERGNVISSHGNPILLNSSIPEDPSIKADINKWRIKLDNY
                     STQELGKTIVYLDGSSQSCRFRECNMGNLICDAMINNNLRHTDEMFWNHVSMCILNGG
                     GIRSPIDERNNGTITWENLAAVLPFGGTFDLVQLKGSTLKKAFEHSVHRYGQSTGEFL
                     QVGGIHVVYDLSRKPGDRVVKLDVLCTKCRVPSYDPLKMDEVYKVILPNFLANGGDGF
                     QMIKDELLRHDSGDQDINVVSTYISKMKVIYPAVEGRIKFSTGSHCHGSFSLIFLSLW
                     AVIFVLYQ"
     misc_feature    170232
                     /note="Clone_right_end: RP11-30P6"
ORIGIN

  Query Match              1.7%;  Score 52;  DB 8;  Length 170232;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128510 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 128459


RESULT 148
AC073068
LOCUS       AC073068               171058 bp    DNA     linear   PRI 08-OCT-2003
DEFINITION  Homo sapiens BAC clone RP11-159I18 from 7, complete sequence.
ACCESSION   AC073068
VERSION     AC073068.8  GI:14010911
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171058)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 171058)
  AUTHORS   Belter,E., Maupin,R. and Podhrasky,A.
  TITLE     The sequence of Homo sapiens BAC clone RP11-159I18
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 171058)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 171058)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 171058)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 171058)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On May 9, 2001 this sequence version replaced gi:10518416.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0159I18
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP11-159I18
            actual end is at base position 171058 of RP11-159I18.
FEATURES                Location/Qualifiers
     source             1. .171058
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP11-159I18"
                     /clone_lib="RPCI-11"
     repeat_region   1. .835
                     /rpt_family="L1"
     repeat_region   839. .1818
                     /rpt_family="L1"
     repeat_region   2300. .2408
                     /rpt_family="ERVL"
     repeat_region   2997. .3302
                     /rpt_family="Alu"
     repeat_region   3483. .3512
                     /rpt_family="AT_rich"
     repeat_region   5688. .5842
                     /rpt_family="L1"
     repeat_region   6268. .6676
                     /rpt_family="L1"
     repeat_region   6693. .7000
                     /rpt_family="Alu"
     repeat_region   7003. .7114
                     /rpt_family="L1"
     repeat_region   7151. .7315
                     /rpt_family="L1"
     repeat_region   7316. .7618
                     /rpt_family="Alu"
     repeat_region   7619. .7838
                     /rpt_family="L1"
     repeat_region   8030. .8166
                     /rpt_family="GA-rich"
     repeat_region   8172. .8986
                     /rpt_family="L1"
     repeat_region   8986. .9561
                     /rpt_family="L1"
     repeat_region   9574. .9723
                     /rpt_family="GA-rich"
     repeat_region   9809. .10484
                     /rpt_family="L1"
     repeat_region   10524. .10552
                     /rpt_family="AT_rich"
     repeat_region   10745. .10968
                     /rpt_family="L1"
     repeat_region   10975. .11399
                     /rpt_family="L1"
     repeat_region   11705. .11773
                     /rpt_family="L1"
     repeat_region   11783. .11985
                     /rpt_family="Alu"
     repeat_region   11986. .12013
                     /rpt_family="L1"
     repeat_region   12014. .12037
                     /rpt_family="(TG)n"
     repeat_region   12038. .13301
                     /rpt_family="L1"
     repeat_region   13302. .13594
                     /rpt_family="Alu"
     repeat_region   13595. .14332
                     /rpt_family="L1"
     repeat_region   14331. .14700
                     /rpt_family="L1"
     repeat_region   16105. .16181
                     /rpt_family="MER1_type"
     repeat_region   17251. .17343
                     /rpt_family="ERV1"
     repeat_region   17360. .17741
                     /rpt_family="ERV1"
     repeat_region   18632. .18662
                     /rpt_family="(TTCC)n"
     repeat_region   18657. .18700
                     /rpt_family="CT-rich"
     repeat_region   19609. .19720
                     /rpt_family="MIR"
     repeat_region   20991. .21033
                     /rpt_family="(TG)n"
     repeat_region   21046. .21083
                     /rpt_family="(TTTTG)n"
     repeat_region   21088. .21271
                     /rpt_family="Alu"
     repeat_region   21754. .21916
                     /rpt_family="MIR"
     repeat_region   22245. .22279
                     /rpt_family="CR1"
     repeat_region   22732. .23580
                     /rpt_family="L1"
     repeat_region   24129. .25030
                     /rpt_family="L1"
     repeat_region   25031. .25336
                     /rpt_family="Alu"
     repeat_region   25337. .25546
                     /rpt_family="L1"
     repeat_region   25550. .25632
                     /rpt_family="Mariner"
     repeat_region   25663. .25997
                     /rpt_family="ERVL"
     repeat_region   26166. .26337
                     /rpt_family="MaLR"
     repeat_region   26338. .26431
                     /rpt_family="ERVL"
     repeat_region   26441. .27629
                     /rpt_family="L1"
     repeat_region   27630. .27683
                     /rpt_family="L1"
     repeat_region   27684. .27970
                     /rpt_family="ERVL"
     repeat_region   27993. .28276
                     /rpt_family="L1"
     repeat_region   28315. .28609
                     /rpt_family="L1"
     repeat_region   29268. .29353

  Query Match             1.7%;  Score 52;  DB 8;  Length 171058;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7545 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7596


RESULT 149
AC010616
LOCUS       AC010616              171849 bp    DNA     linear   PRI 01-MAY-2001
DEFINITION  Homo sapiens chromosome 19 clone CTD-2575K13, complete sequence.
ACCESSION   AC010616
VERSION     AC010616.5  GI:13876489
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171849)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 171849)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 171849)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 171849)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 28, 2001 this sequence version replaced gi:7711477.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 1.
            STS Content:
            WI-15934 G21486
            SHGC-64254 G38544.
FEATURES             Location/Qualifiers
     source          1. .171849
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2575K13"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 171849;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129747 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 129798


RESULT 150
AL590609/c
LOCUS       AL590609              172307 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-4M23 on chromosome 1 Contains
            the 5' end of the FUCA1 gene for alpha-L- 1 fucosidase (tissue),
            the CNR2 gene for cannabinoid receptor 2 (macrophage), a BTB (POZ)
            domain-containing 6 (BTBD6) pseudogene, a novel gene, an H3 histone
            family 3A (H3F3A) pseudogene, the PNRC2 gene for proline-rich
            nuclear receptor coactivator 2, the FUS1P1 gene for FUS interacting
            protein (serine-arginine rich) 1, a ribosomal protein L36 (RPL36)
            pseudogene and seven CpG islands, complete sequence.
ACCESSION   AL590609
VERSION     AL590609.15  GI:39573520
KEYWORDS    HTG; BTBD6; cannabinoid receptor; CNR2; CpG island; FUC1A; FUS;
            FUS1P1; H3F3A; histone; PNCR2; POZ; proline-rich; RPL36.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 172307)
  AUTHORS   Wallis,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 8, 2003 this sequence version replaced gi:15795484.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-4M23 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .172307
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-4M23"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-4M23"
     gene            complement(join(3548. .8684,46268. .46384))
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
     mRNA            complement(join(3548. .8684,46268. .46384))
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
                     /product="cannabinoid receptor 2 (macrophage)"
                     /note="match: ESTs: Em:AA485015.1 Em:AA909480.1
                     Em:AI218261.1 Em:AI424994.1 Em:AI475798.1 Em:AI864752.1
                     Em:AW503816.1 Em:BM151899.1
                     match: cDNAs: Em:X74328.1"
     polyA_site      complement(3548)
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
     polyA_signal    complement(3566. .3571)
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
     polyA_site      complement(6097)
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
     polyA_signal    complement(6116. .6121)
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
     CDS             complement(7557. .8639)
                     /gene="CNR2"
                     /locus_tag="RP11-4M23.1-001"
                     /standard_name="OTTHUMP00000015772"
                     /note="match: proteins: Sw:P34972 Sw:P47936 Sw:Q9QZN9
                     Tr:BAC29520 Tr:BAC29894"
                     /codon_start=1
                     /product="cannabinoid receptor 2 (macrophage)"
                     /protein_id="CAI14799.1"
                     /db_xref="GI:55961036"
                     /db_xref="InterPro:IPR000276"
                     /db_xref="InterPro:IPR001551"
                     /db_xref="InterPro:IPR002230"
                     /translation="MEECWVTEIANGSKDGLDSNPMKDYMILSGPQKTAVAVLCTLLG
                     LLSALENVAVLYLILSSHQLRRKPSYLFIGSLAGADFLASVVFACSFVNFHVFHGVDS
                     KAVFLLKIGSVTMTFTASVGSLLLTAIDRYLCLRYPPSYKALLTRGRALVTLGIMWVL
                     SALVSYLPLMGWTCCPRPCSELFPLIPNDYLLSWLLFIAFLFSGIIYTYGHVLWKAHQ
                     HVASLSGHQDRQVPGMARMRLDVRLAKTLGLVLAVLLICWFPVLALMAHSLATTLSDQ
                     VKKAFAFCSMLCLINSMVNPVIYALRSGEIRSSAHHCLAHWKKCVRGLGSEAKEEAPR
                     SSVTETEADGKITPWPDSRDLDLSDC"
     gene            complement(34493. .35759)
                     /locus_tag="RP11-4M23.2-001"
                     /pseudo
     CDS             complement(34493. .35759)
                     /locus_tag="RP11-4M23.2-001"
```

```
                     /note="match: proteins: Sw:P58545 Sw:Q96KE9 Sw:Q9Y2F9
                     Tr:AAH62968 Tr:Q8CHC1 Tr:Q9NST6"
                     /pseudo
                     /codon_start=1
                     /product="BTB (POZ) domain-containing 6 (BTBD6)
                     pseudogene"
     gene            complement(join(40133. .40203,40769. .40907))
                     /locus_tag="RP11-4M23.3-001"
     mRNA            complement(join(40133. .40203,40769. .40907))
                     /locus_tag="RP11-4M23.3-001"
                     /product="putative novel transcript"
                     /note="match: ESTs: Em:CD696420.1"
     misc_feature    73316
                     /note="Clone_right_end: RP11-45G17"
     gene            complement(82037. .82405)
                     /locus_tag="RP11-4M23.4-001"
                     /pseudo
     CDS             complement(82037. .82405)
                     /locus_tag="RP11-4M23.4-001"
                     /note="match: proteins: Sw:Q92068 Tr:AAH12687 Tr:AAH17558
                     Tr:BAB22464 Tr:BAC29895 Tr:BAC40130 Tr:Q7ZV67 Tr:Q8WSF1
                     Tr:Q9U281"
                     /pseudo
                     /codon_start=1
                     /product="H3 histone, family 3A (H3F3A) pseudogene"
     gene            join(92131. .92278,93951. .94156,94462. .94487)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-003"
     mRNA            join(92131. .92278,93951. .94156,94462. .94487)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-003"
                     /product="proline-rich nuclear receptor coactivator 2"
                     /note="match: ESTs: Em:AV686572.1 Em:CD689388.1"
     gene            join(92834. .93020,93951. .94156,94462. .96484)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     mRNA            join(92834. .93020,93951. .94156,94462. .96484)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
                     /product="proline-rich nuclear receptor coactivator 2"
                     /note="match: ESTs: Em:AI221857.1 Em:AI240994.1
                     Em:AI684547.1 Em:AI823428.1 Em:AI829468.1 Em:AU280480.1
                     Em:AW184027.1 Em:AW274263.1 Em:BE047281.1 Em:BM922437.1
                     Em:BU159988.1 Em:BU165438.1
                     match: cDNAs: Em:AF151042.1 Em:AF374386.1 Em:AK000319.1
                     Em:AK024774.1 Em:BC001959.1"
     gene            join(92850. .93020,93955. .94156,94462. .94902)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-002"
     mRNA            join(92850. .93020,93955. .94156,94462. .94902)
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-002"
                     /product="proline-rich nuclear receptor coactivator 2"
                     /note="match: ESTs: Em:BG715201.1 Em:BI596502.1"
     CDS             94480. .94899
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
                     /standard_name="OTTHUMP00000015770"
                     /note="match: proteins: Tr:Q9CR73 Tr:Q9CXC6 Tr:Q9NPJ4"
                     /codon_start=1
                     /product="proline-rich nuclear receptor coactivator 2"
                     /protein_id="CAI14802.1"
                     /db_xref="GI:55961037"
                     /db_xref="GOA:Q9NPJ4"
                     /db_xref="UniProt/Swiss-Prot:Q9NPJ4"
                     /translation="MGGGERYNIPAPQSRNVSKNQQQLNRQKTKEQNSQMKIVHKKKE
                     RGHGYNSSAAAWQAMQNGGKNKNFPNNQSWNSSLSGPRLLFKSQANQNYAGAKFSEPP
                     SPSVLPKPPSHWVPVSFNPSDKEIMTFQLKTLLKVQV"
     polyA_signal    95681. .95686
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     polyA_site      95703
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     polyA_signal    96441. .96446
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     polyA_signal    96445. .96450
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     polyA_site      96484
                     /gene="PNRC2"
                     /locus_tag="RP11-4M23.5-001"
     gene            complement(join(97826. .100745,104595. .104648,
                     104872. .105034,107996. .108099,111735. .111839,
                     113213. .113378))
                     /gene="FUSIP1"
                     /locus_tag="RP11-4M23.6-002"
     mRNA            complement(join(97826. .100745,104595. .104648,
                     104872. .105034,107996. .108099,111735. .111839,
                     113213. .113378))

  Query Match             1.7%;  Score 52;  DB 8;  Length 172307;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     149550 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 149499


RESULT 151
AC025589
LOCUS       AC025589              172759 bp    DNA     linear   PRI 01-MAY-2001
DEFINITION  Homo sapiens chromosome 3 clone RP11-175F9, complete sequence.
ACCESSION   AC025589
VERSION     AC025589.20  GI:13899418
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 172759)
  AUTHORS   Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
            Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
            Mao,J., Komp,C., Kottler,S., Lam,B., Marathe,R., Miranda,M.,
            Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
            Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172759)
  AUTHORS   Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D., Federspiel,N.,
            Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
            Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,
            Yu,S. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 172759)
  AUTHORS   Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
            Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
            Komp,C., Kottler,S., Lam,B., Mao,J., Marathe,R., Miranda,M.,
            Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
            Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2001) Genome Technology Center, Stanford
            University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT     On May 1, 2001 this sequence version replaced gi:13446262.
            -------------- Genome Center
                 Center: Stanford DNA Sequencing and Technology Development
            Center
                 Center code: SDSTDC
                 Web site: http://sequence-www.stanford.edu/group/human/
                 Contact: hum-info@sequence.stanford.edu
            -------------- Project Information
```

```
                Center project name: 860
                Center clone name: RP11-175F9

             ---------- Summary
             This sequence was finished as follows unless otherwise noted:
             all regions were double stranded or sequenced with an alternate
             chemistry or covered by high quality data (i.e., phred quality >=
             30); an attempt was made to resolve all sequencing problems, such
             as compressions and repeats; all regions were  covered by sequence
             from more than one subclone; and the assembly was confirmed by
             restriction digest.
FEATURES              Location/Qualifiers
     source           1. .172759
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="3"
                      /clone="RP11-175F9"
                      /clone_lib="RPCI human BAC library 11"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 172759;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14184 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 14235


RESULT 152
AC010160/c
LOCUS       AC010160              172915 bp    DNA     linear   HTG 01-OCT-2002
DEFINITION  Homo sapiens chromosome 10 clone RP11-396M20, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC010160
VERSION     AC010160.9  GI:13958383
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN;
            HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 172915)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172915)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On May 6, 2001 this sequence version replaced gi:9887635.
            -------------- Genome Center
                           Center: Genome Therapeutics Corporation
                           Center code: GTC
                           Web site: http://www.genomecorp.com/
                           Contact: gtc-seqcenter@genomecorp.com
                        -------------- Project Information
                           Center project name: hg015
                        -------------- Summary Statistics
                           Sequencing vector: N/A
                           Chemistry: Dye-terminator Big Dye; 100% of reads
                           Assembly program: Phrap; version 990315
                           Consensus quality: 166031 bases at least Q40
                           Consensus quality: 167418 bases at least Q30
                           Consensus quality: 168334 bases at least Q20
                           Insert size: 172565; sum-of-contigs
                           Quality coverage: 7.4x in Q20 bases; sum-of-contigs
                        --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1479: contig of 1479 bp in length
            *     1480     1579: gap of unknown length
            *     1580    19522: contig of 17943 bp in length
            *    19523    19622: gap of unknown length
            *    19623    43164: contig of 23542 bp in length
            *    43165    43264: gap of unknown length
            *    43265    88435: contig of 45171 bp in length
            *    88436    88535: gap of unknown length
            *    88536   172915: contig of 84380 bp in length.
FEATURES              Location/Qualifiers
     source           1. .172915
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="10"
                      /clone="RP11-396M20"
                      /clone_lib="RPCI-11"
     misc_feature     1. .1479
                      /note="assembly_name:Contig3"
     gap              1480. .1579
                      /estimated_length=unknown
     misc_feature     1580. .19522
                      /note="assembly_name:Contig5"
     gap              19523. .19622
                      /estimated_length=unknown
     misc_feature     19623. .43164
                      /note="assembly_name:Contig6"
     gap              43165. .43264
                      /estimated_length=unknown
     misc_feature     43265. .88435
                      /note="assembly_name:Contig7"
     gap              88436. .88535
                      /estimated_length=unknown
     misc_feature     88536. .172915
                      /note="assembly_name:Contig8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 172915;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153187 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 153136


RESULT 153
AC020908
LOCUS       AC020908              174034 bp    DNA     linear   PRI 03-OCT-2001
DEFINITION  Homo sapiens chromosome 19 clone CTD-2528A14, complete sequence.
ACCESSION   AC020908
VERSION     AC020908.7  GI:15887300
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 174034)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174034)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
```

```
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 174034)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 174034)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Oct 3, 2001 this sequence version replaced gi:9958017.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.6.
FEATURES             Location/Qualifiers
     source          1. .174034
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2528A14"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 174034;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61038 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 61089


RESULT 154
AC127468/c
LOCUS       AC127468               174521 bp    DNA     linear   HTG 05-JUN-2003
DEFINITION  Papio anubis clone RP41-263F8, WORKING DRAFT SEQUENCE, 15 ordered
            pieces.
ACCESSION   AC127468
VERSION     AC127468.4  GI:31415891
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 174521)
  AUTHORS   Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
            Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
            Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
            Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
            Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
            Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
            Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
            Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
            Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174521)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 174521)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2003) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Jun 5, 2003 this sequence version replaced gi:26050953.
            ------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            ------------- Project Information
                 Center project name: cyr
                 Center clone name: 263F08

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            ------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 171309 bases at least Q40
                 Consensus quality: 172311 bases at least Q30
                 Consensus quality: 172845 bases at least Q20
                 Insert size: 177000; agarose-fp
                 Insert size: 173121; sum-of-contigs
                 Quality coverage: 10.09x in Q20 bases; agarose-fp
                 Quality coverage: 10.32x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1    42198: contig of 42198 bp in length
            *   42199    42298: gap of unknown length
            *   42299    45686: contig of 3388 bp in length
            *   45687    45786: gap of unknown length
            *   45787    71568: contig of 25782 bp in length
            *   71569    71668: gap of unknown length
            *   71669    75986: contig of 4318 bp in length
            *   75987    76086: gap of unknown length
            *   76087    85283: contig of 9197 bp in length
            *   85284    85383: gap of unknown length
            *   85384   103813: contig of 18430 bp in length
            *  103814   103913: gap of unknown length
            *  103914   105217: contig of 1304 bp in length
            *  105218   105317: gap of unknown length
            *  105318   112563: contig of 7246 bp in length
            *  112564   112663: gap of unknown length
            *  112664   114712: contig of 2049 bp in length
            *  114713   114812: gap of unknown length
            *  114813   141386: contig of 26574 bp in length
            *  141387   141486: gap of unknown length
            *  141487   144449: contig of 2963 bp in length
            *  144450   144549: gap of unknown length
            *  144550   148916: contig of 4367 bp in length
            *  148917   149016: gap of unknown length
            *  149017   153247: contig of 4231 bp in length
            *  153248   153347: gap of unknown length
            *  153348   165212: contig of 11865 bp in length
            *  165213   165312: gap of unknown length
            *  165313   174521: contig of 9209 bp in length.
FEATURES             Location/Qualifiers
     source          1. .174521
```

```
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="RP41-263F8"
                     /clone_lib="RP41"
     misc_feature    1. .42198
                     /note="assembly_fragment
                     missing T7 clone end on 5' end of insert"
     gap             42199. .42298
                     /estimated_length=unknown
     misc_feature    42299. .45686
                     /note="assembly_fragment"
     gap             45687. .45786
                     /estimated_length=unknown
     misc_feature    45787. .71568
                     /note="assembly_fragment"
     gap             71569. .71668
                     /estimated_length=unknown
     misc_feature    71669. .75986
                     /note="assembly_fragment"
     gap             75987. .76086
                     /estimated_length=unknown
     misc_feature    76087. .85283
                     /note="assembly_fragment"
     gap             85284. .85383
                     /estimated_length=unknown
     misc_feature    85384. .103813
                     /note="assembly_fragment"
     gap             103814. .103913
                     /estimated_length=unknown
     misc_feature    103914. .105217
                     /note="assembly_fragment"
     gap             105218. .105317
                     /estimated_length=unknown
     misc_feature    105318. .174521
                     /note="clone overlaps with GenBank Accession Number
                     AC127469 clone RP41-91L23 (center project name cyq)"
     misc_feature    105318. .112563
                     /note="assembly_fragment"
     gap             112564. .112663
                     /estimated_length=unknown
     misc_feature    112664. .114712
                     /note="assembly_fragment"
     gap             114713. .114812
                     /estimated_length=unknown
     misc_feature    114813. .141386
                     /note="assembly_fragment"
     gap             141387. .141486
                     /estimated_length=unknown
     misc_feature    141487. .144449
                     /note="assembly_fragment"
     gap             144450. .144549
                     /estimated_length=unknown
     misc_feature    144550. .148916
                     /note="assembly_fragment"
     gap             148917. .149016
                     /estimated_length=unknown
     misc_feature    149017. .153247
                     /note="assembly_fragment"
     gap             153248. .153347
                     /estimated_length=unknown
     misc_feature    153348. .165212
                     /note="assembly_fragment"
     gap             165213. .165312
                     /estimated_length=unknown
     misc_feature    165313. .174521
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 174521;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;

  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90300 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 90249


RESULT 155
CNS01DX3/c
LOCUS       CNS01DX3               176343 bp    DNA     linear   PRI 26-APR-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-517O13 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL139021
VERSION     AL139021.6  GI:13159194
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176343)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 176343)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Feb 27, 2001 this sequence version replaced gi:11611151.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : C-2002H8
            Downstream BAC (overlapping the SP6 end) : R-401H9 (AC=AL135752)
            -------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 8.11x in Q20 bases; sum-of-contigs
            --------------
            ----------------------
            Overall quality chart :
            Range     :   bases
            0         :
            1  -  9   :
            10 - 19   :       2
            20 - 29   :      10
            30 - 39   :     210
            40 - 49   :    4379
            50 - 59   :   10906
            60 - 69   :    9685
            70 - 79   :   18251
            80 - 89   :   45109
            90 - 99   :   87791
            ----------------------
            Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1. .176343
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-517O13"
                     /clone_lib="RPCI-11"
     STS             2338. .2477
                     /note="matching EMBL:G33047
```

```
                     RHdb:RH67742
                     dbSTS:STS47669
                     Identified using the e-PCR software (G. Schuler)"
     STS             35745. .35866
                     /note="matching EMBL:G13592
                     RHdb:RH95857
                     RHdb:RH13847
                     dbSTS:STS13237
                     Identified using the e-PCR software (G. Schuler)"
     STS             37127. .37251
                     /note="matching EMBL:G30370
                     RHdb:RH53945
                     RHdb:RH38563
                     dbSTS:STS8453
                     Identified using the e-PCR software (G. Schuler)"
     STS             66498. .66635
                     /note="matching EMBL:N55165
                     RHdb:RH66721
                     dbSTS:STS46650
                     Identified using the e-PCR software (G. Schuler)"
     STS             71248. .71368
                     /note="matching EMBL:H97583
                     RHdb:RH78893
                     dbSTS:STS55634
                     Identified using the e-PCR software (G. Schuler)"
     STS             72428. .72563
                     /note="matching EMBL:N22461
                     RHdb:RH46962
                     dbSTS:STS40024
                     Identified using the e-PCR software (G. Schuler)"
     STS             110431. .110732
                     /note="matching EMBL:G03433
                     RHdb:RH34713
                     dbSTS:STS1533
                     Identified using the e-PCR software (G. Schuler)"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 176343;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4072 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4021


RESULT 156
AC150824
LOCUS       AC150824               176871 bp    DNA     linear   HTG 16-SEP-2004
DEFINITION  Callithrix jacchus clone CH259-392G13, WORKING DRAFT SEQUENCE, 6
            ordered pieces.
ACCESSION   AC150824
VERSION     AC150824.2  GI:52138777
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Callithrix jacchus (white-tufted-ear marmoset)
  ORGANISM  Callithrix jacchus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Callitrichidae; Callithrix.
REFERENCE   1  (bases 1 to 176871)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H., Daki,N.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Park,M., Portnoy,M.E., Prasad,A., Puri,O.,
            Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G., Shah,K.,
            Sison,C., Stantripop,S., Stephen,E., Thomas,J.W., Thomas,P.J.,
            Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 176871)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 176871)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Sep 16, 2004 this sequence version replaced gi:51101028.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: hjb
                 Center clone name: 392G13

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 175488 bases at least Q40
                 Consensus quality: 176005 bases at least Q30
                 Consensus quality: 176238 bases at least Q20
                 Insert size: 225000; agarose-fp
                 Insert size: 176371; sum-of-contigs
                 Quality coverage: 8.55x in Q20 bases; agarose-fp
                 Quality coverage: 10.91x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    22851: contig of 22851 bp in length
            *    22852    22951: gap of unknown length
            *    22952    25041: contig of 2090 bp in length
            *    25042    25141: gap of unknown length
            *    25142    63108: contig of 37967 bp in length
            *    63109    63208: gap of unknown length
            *    63209   156929: contig of 93721 bp in length
            *   156930   157029: gap of unknown length
            *   157030   171090: contig of 14061 bp in length
            *   171091   171190: gap of unknown length
            *   171191   176871: contig of 5681 bp in length.
FEATURES             Location/Qualifiers
     source          1. .176871
                     /organism="Callithrix jacchus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9483"
                     /clone="CH259-392G13"
                     /clone_lib="CH259"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .106470
                     /note="clone overlaps with GenBank Accession Number
```

```
                AC151042 clone CH259-63D17 (center project name hja)"
     misc_feature    1. .22851
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     gap             22852. .22951
                     /estimated_length=unknown
     misc_feature    22952. .25041
                     /note="assembly_fragment"
     gap             25042. .25141
                     /estimated_length=unknown
     misc_feature    25142. .63108
                     /note="assembly_fragment"
     gap             63109. .63208
                     /estimated_length=unknown
     misc_feature    63209. .156929
                     /note="assembly_fragment"
     misc_feature    79212. .176871
                     /note="clone overlaps with GenBank Accession Number
                     AC151035 clone CH259-434P22 (center project name hjc)"
     gap             156930. .157029
                     /estimated_length=unknown
     misc_feature    157030. .171090
                     /note="assembly_fragment"
     gap             171091. .171190
                     /estimated_length=unknown
     misc_feature    171191. .176871
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 176871;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4700 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4751


RESULT 157
AC073954/c
LOCUS       AC073954               177744 bp    DNA     linear   HTG 17-NOV-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
            SEQUENCE, 34 unordered pieces.
ACCESSION   AC073954
VERSION     AC073954.3  GI:10998993
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 177744)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177744)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUL-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Oct 25, 2000 this sequence version replaced gi:9186849.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0147C10
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 155679 bases at least Q40
            Consensus quality: 163329 bases at least Q30
            Consensus quality: 166780 bases at least Q20
            Insert size: 185000; agarose-fp
            Insert size: 183559; sum-of-contigs
            Quality coverage: 3.29 in Q20 bases; agarose-fp
            Quality coverage: 3.43 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 34 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1247: contig of 1247 bp in length
            *    1248     1347: gap of unknown length
            *    1348     2762: contig of 1415 bp in length
            *    2763     2862: gap of unknown length
            *    2863     4773: contig of 1911 bp in length
            *    4774     4873: gap of unknown length
            *    4874     6326: contig of 1453 bp in length
            *    6327     6426: gap of unknown length
            *    6427     8873: contig of 2447 bp in length
            *    8874     8973: gap of unknown length
            *    8974    10783: contig of 1810 bp in length
            *   10784    10883: gap of unknown length
            *   10884    12763: contig of 1880 bp in length
            *   12764    12863: gap of unknown length
            *   12864    15381: contig of 2518 bp in length
            *   15382    15481: gap of unknown length
            *   15482    17365: contig of 1884 bp in length
            *   17366    17465: gap of unknown length
            *   17466    19992: contig of 2527 bp in length
            *   19993    20092: gap of unknown length
            *   20093    23512: contig of 3420 bp in length
            *   23513    23612: gap of unknown length
            *   23613    25699: contig of 2087 bp in length
            *   25700    25799: gap of unknown length
            *   25800    28333: contig of 2534 bp in length
            *   28334    28433: gap of unknown length
            *   28434    32027: contig of 3594 bp in length
            *   32028    32127: gap of unknown length
            *   32128    36115: contig of 3988 bp in length
            *   36116    36215: gap of unknown length
            *   36216    39882: contig of 3667 bp in length
            *   39883    39982: gap of unknown length
            *   39983    44009: contig of 4027 bp in length
            *   44010    44109: gap of unknown length
            *   44110    47829: contig of 3720 bp in length
            *   47830    47929: gap of unknown length
            *   47930    53768: contig of 5839 bp in length
            *   53769    53868: gap of unknown length
            *   53869    60215: contig of 6347 bp in length
            *   60216    60315: gap of unknown length
            *   60316    66477: contig of 6162 bp in length
            *   66478    66577: gap of unknown length
            *   66578    71476: contig of 4899 bp in length
            *   71477    71576: gap of unknown length
            *   71577    78883: contig of 7307 bp in length
            *   78884    78983: gap of unknown length
            *   78984    85952: contig of 6969 bp in length
            *   85953    86052: gap of unknown length
            *   86053    93510: contig of 7458 bp in length
            *   93511    93610: gap of unknown length
            *   93611   100718: contig of 7108 bp in length
            *  100719   100818: gap of unknown length
```

```
            *   100819   109795: contig of 8977 bp in length
            *   109796   109895: gap of unknown length
            *   109896   117790: contig of 7895 bp in length
            *   117791   117890: gap of unknown length
            *   117891   126044: contig of 8154 bp in length
            *   126045   126144: gap of unknown length
            *   126145   134442: contig of 8298 bp in length
            *   134443   134542: gap of unknown length
            *   134543   145663: contig of 11121 bp in length
            *   145664   145763: gap of unknown length
            *   145764   157189: contig of 11426 bp in length
            *   157190   157289: gap of unknown length
            *   157290   166822: contig of 9533 bp in length
            *   166823   166922: gap of unknown length
            *   166923   177744: contig of 10822 bp in length.
FEATURES             Location/Qualifiers
     source          1. .177744
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="RP11-147C10"
     misc_feature    1. .1247
                     /note="assembly_name:Contig26"
     gap             1248. .1347
                     /estimated_length=unknown
     misc_feature    1348. .2762
                     /note="assembly_name:Contig31"
     gap             2763. .2862
                     /estimated_length=unknown
     misc_feature    2863. .4773
                     /note="assembly_name:Contig34"
     gap             4774. .4873
                     /estimated_length=unknown
     misc_feature    4874. .6326
                     /note="assembly_name:Contig35"
     gap             6327. .6426
                     /estimated_length=unknown
     misc_feature    6427. .8873
                     /note="assembly_name:Contig36"
     gap             8874. .8973
                     /estimated_length=unknown
     misc_feature    8974. .10783
                     /note="assembly_name:Contig37"
     gap             10784. .10883
                     /estimated_length=unknown
     misc_feature    10884. .12763
                     /note="assembly_name:Contig38"
     gap             12764. .12863
                     /estimated_length=unknown
     misc_feature    12864. .15381
                     /note="assembly_name:Contig39"
     gap             15382. .15481
                     /estimated_length=unknown
     misc_feature    15482. .17365
                     /note="assembly_name:Contig40"
     gap             17366. .17465
                     /estimated_length=unknown
     misc_feature    17466. .19992
                     /note="assembly_name:Contig41"
     gap             19993. .20092
                     /estimated_length=unknown
     misc_feature    20093. .23512
                     /note="assembly_name:Contig42"
     gap             23513. .23612
                     /estimated_length=unknown
     misc_feature    23613. .25699
                     /note="assembly_name:Contig43"
     gap             25700. .25799
                     /estimated_length=unknown
     misc_feature    25800. .28333
                     /note="assembly_name:Contig44"
     gap             28334. .28433
                     /estimated_length=unknown
     misc_feature    28434. .32027
                     /note="assembly_name:Contig45"
     gap             32028. .32127
                     /estimated_length=unknown
     misc_feature    32128. .36115
                     /note="assembly_name:Contig46"
     gap             36116. .36215
                     /estimated_length=unknown
     misc_feature    36216. .39882
                     /note="assembly_name:Contig47"
     gap             39883. .39982
                     /estimated_length=unknown
     misc_feature    39983. .44009
                     /note="assembly_name:Contig48"
     gap             44010. .44109
                     /estimated_length=unknown
     misc_feature    44110. .47829
                     /note="assembly_name:Contig49"
     gap             47830. .47929
                     /estimated_length=unknown
     misc_feature    47930. .53768
                     /note="assembly_name:Contig50"
     gap             53769. .53868

  Query Match             1.7%;  Score 52;  DB 14;  Length 177744;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146170 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 146119


RESULT 158
AC010761
LOCUS       AC010761              177773 bp    DNA     linear   PRI 10-JAN-2003
DEFINITION  Homo sapiens chromosome 17, clone RP11-386F9, complete sequence.
ACCESSION   AC010761
VERSION     AC010761.10  GI:27573488
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 177773)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-386F9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177773)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 177773)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
```

```
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 177773)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 10, 2003 this sequence version replaced gi:24850532.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1611
                Center clone name: 386_F_9
            --------------.
FEATURES             Location/Qualifiers
     source          1. .177773
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-386F9"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   31. .91
                     /rpt_family="MIR"
     repeat_region   92. .396
                     /rpt_family="AluY"
     repeat_region   397. .443
                     /rpt_family="MIR"
     repeat_region   1541. .1662
                     /rpt_family="AluJb"
     repeat_region   1663. .1915
                     /rpt_family="AluSp"
     repeat_region   1916. .2068
                     /rpt_family="AluJb"
     repeat_region   complement(2094. .2403)
                     /rpt_family="AluSp"
     repeat_region   complement(2621. .2655)
                     /rpt_family="FAM"
     repeat_region   2656. .2983
                     /rpt_family="AluSq"
     repeat_region   complement(2984. .3026)
                     /rpt_family="FAM"
     repeat_region   complement(3028. .3215)
                     /rpt_family="AluJo"
     repeat_region   complement(3230. .3515)
                     /rpt_family="AluSx"
     repeat_region   3607. .3773
                     /rpt_family="MIR"
     repeat_region   3932. .3953
                     /rpt_family="(TG)n"
     repeat_region   3956. .4172
                     /rpt_family="L1MC3"
     repeat_region   4177. .4349
                     /rpt_family="FAM"
     repeat_region   complement(4456. .4759)
                     /rpt_family="AluSg"
     repeat_region   4772. .5682
                     /rpt_family="L1MC3"
     repeat_region   5683. .6006
                     /rpt_family="AluSg"
     repeat_region   6007. .6108
                     /rpt_family="L1MC3"
     repeat_region   6181. .6455
                     /rpt_family="AluJo"
     repeat_region   6685. .6821
                     /rpt_family="AluSx"
     repeat_region   6822. .7132
                     /rpt_family="AluSg"
     repeat_region   7133. .7294
                     /rpt_family="AluSx"
     repeat_region   8303. .8367
                     /rpt_family="Alu"
     repeat_region   complement(8373. .8421)
                     /rpt_family="Alu"
     repeat_region   8431. .8455
                     /rpt_family="AT_rich"
     repeat_region   8456. .8759
                     /rpt_family="AluSx"
     repeat_region   9104. .9406
                     /rpt_family="AluSg"
     repeat_region   9419. .9536
                     /rpt_family="FLAM_C"
     repeat_region   complement(9557. .9849)
                     /rpt_family="AluJb"
     repeat_region   9850. .9883
                     /rpt_family="AT_rich"
     repeat_region   9886. .10131
                     /rpt_family="AluJb"
     repeat_region   complement(10405. .11595)
                     /rpt_family="SVA"
     repeat_region   complement(11607. .12423)
                     /rpt_family="SVA"
     repeat_region   complement(12650. .12763)
                     /rpt_family="MER5A"
     repeat_region   13119. .13194
                     /rpt_family="L2"
     repeat_region   13236. .13255
                     /rpt_family="(A)n"
     repeat_region   13491. .13699
                     /rpt_family="MIR"
     repeat_region   complement(15981. .16105)
                     /rpt_family="FRAM"
     repeat_region   16137. .16377
                     /rpt_family="L2"
     repeat_region   16378. .16660
                     /rpt_family="AluSx"
     repeat_region   16661. .17274
```

```
                     /rpt_family="L2"
     repeat_region   17513. .17625
                     /rpt_family="L2"
     repeat_region   17629. .17922
                     /rpt_family="AluSx"
     repeat_region   17924. .17949
                     /rpt_family="(TATATG)n"
     repeat_region   17950. .18079
                     /rpt_family="FRAM"
     repeat_region   18097. .18145
                     /rpt_family="L2"
     repeat_region   complement(18703. .19008)
                     /rpt_family="AluSx"
     repeat_region   19259. .19546
                     /rpt_family="AluY"
     repeat_region   19558. .19586
                     /rpt_family="(TA)n"

  Query Match             1.7%;  Score 52;  DB 8;  Length 177773;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45101 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 45152


RESULT 159
AC012350/c
LOCUS       AC012350              178168 bp    DNA     linear   HTG 01-APR-2000
DEFINITION  Homo sapiens clone RP11-16N9, WORKING DRAFT SEQUENCE, 16 unordered
            pieces.
ACCESSION   AC012350
VERSION     AC012350.3  GI:7381803
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 178168)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-16N9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 178168)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 1, 2000 this sequence version replaced gi:6479001.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3593
                Center clone name: 16_N_9
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 172070 bases at least Q40
                Consensus quality: 174479 bases at least Q30
                Consensus quality: 175597 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 176668; sum-of-contigs
                Quality coverage: 6.4 in Q20 bases; agarose-fp
                Quality coverage: 6.2 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1694: contig of 1694 bp in length
            *     1695     1794: gap of 100 bp
            *     1795     4437: contig of 2643 bp in length
            *     4438     4537: gap of 100 bp
            *     4538     6644: contig of 2107 bp in length
            *     6645     6744: gap of 100 bp
            *     6745    10092: contig of 3348 bp in length
            *    10093    10192: gap of 100 bp
            *    10193    11801: contig of 1609 bp in length
            *    11802    11901: gap of 100 bp
            *    11902    15584: contig of 3683 bp in length
            *    15585    15684: gap of 100 bp
            *    15685    19675: contig of 3991 bp in length
            *    19676    19775: gap of 100 bp
            *    19776    27913: contig of 8138 bp in length
            *    27914    28013: gap of 100 bp
            *    28014    34156: contig of 6143 bp in length
            *    34157    34256: gap of 100 bp
            *    34257    43325: contig of 9069 bp in length
            *    43326    43425: gap of 100 bp
            *    43426    55265: contig of 11840 bp in length
            *    55266    55365: gap of 100 bp
            *    55366    68306: contig of 12941 bp in length
            *    68307    68406: gap of 100 bp
            *    68407    82910: contig of 14504 bp in length
            *    82911    83010: gap of 100 bp
            *    83011   102181: contig of 19171 bp in length
            *   102182   102281: gap of 100 bp
            *   102282   128058: contig of 25777 bp in length
            *   128059   128158: gap of 100 bp
            *   128159   178168: contig of 50010 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178168
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-16N9"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1694
                     /note="assembly_fragment"
     gap             1695. .1794
                     /estimated_length=100
     misc_feature    1795. .4437
                     /note="assembly_fragment"
     gap             4438. .4537
                     /estimated_length=100
     misc_feature    4538. .6644
                     /note="assembly_fragment"
     gap             6645. .6744
                     /estimated_length=100
     misc_feature    6745. .10092
                     /note="assembly_fragment"
```

```
     gap             10093. .10192
                     /estimated_length=100
     misc_feature    10193. .11801
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             11802. .11901
                     /estimated_length=100
     misc_feature    11902. .15584
                     /note="assembly_fragment"
     gap             15585. .15684
                     /estimated_length=100
     misc_feature    15685. .19675
                     /note="assembly_fragment"
     gap             19676. .19775
                     /estimated_length=100
     misc_feature    19776. .27913
                     /note="assembly_fragment"
     gap             27914. .28013
                     /estimated_length=100
     misc_feature    28014. .34156
                     /note="assembly_fragment"
     gap             34157. .34256
                     /estimated_length=100
     misc_feature    34257. .43325
                     /note="assembly_fragment"
     gap             43326. .43425
                     /estimated_length=100
     misc_feature    43426. .55265
                     /note="assembly_fragment"
     gap             55266. .55365
                     /estimated_length=100
     misc_feature    55366. .68306
                     /note="assembly_fragment"
     gap             68307. .68406
                     /estimated_length=100
     misc_feature    68407. .82910
                     /note="assembly_fragment"
     gap             82911. .83010
                     /estimated_length=100
     misc_feature    83011. .102181
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             102182. .102281
                     /estimated_length=100
     misc_feature    102282. .128058
                     /note="assembly_fragment"
     gap             128059. .128158
                     /estimated_length=100
     misc_feature    128159. .178168
                     /note="assembly_fragment"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 178168;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98028 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 97977

RESULT 160
AC093126/c
LOCUS       AC093126              181792 bp    DNA     linear   PRI 19-OCT-2002
DEFINITION  Papio anubis clone RP41-231J21, complete sequence.
ACCESSION   AC093126
VERSION     AC093126.2  GI:24137422
KEYWORDS    HTG.
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 181792)
  AUTHORS   Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Cariaga,K., Coleman,B., Dietrich,N.L., Engle,J., Granite,S.,
            Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L.,
            Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q., Legaspi,R.,
            Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B.,
            Mastrian,S.D., McCloskey,J.C., McDowell,J., Paguirigan,C.,
            Pearson,R., Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandler,K.,
            Schueler,M.G., Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J.,
            Touchman,J.W., Vogt,J.L., Walker,M., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 181792)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 181792)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Oct 19, 2002 this sequence version replaced gi:15148116.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ccu
                 Center clone name: 231J21

            This sequence was finished as follows unless otherwise noted:
            all regions were double-stranded, sequenced with an
            alternate chemistry, or covered by high quality data
            (i.e., phred quality >= 30); an attempt was made to resolve
            all sequencing problems, such as compressions and repeats;
            all regions were covered by at least one plasmid subclone
            or more than one M13 subclone; and the assembly was confirmed
            by restriction digest.

            CLONE LENGTH: This sequence represents the entire insert of
            this clone unless otherwise noted. If there are overlapping
            clones, the overlaps are noted in the beginning and end of
            the Features section.
FEATURES             Location/Qualifiers
     source          1. .181792
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="RP41-231J21"
                     /clone_lib="RP41"
     misc_feature    19955. .20051
                     /note="single clone coverage"
     misc_feature    44150. .44183
                     /note="single clone coverage"
     misc_feature    55640. .55715
                     /note="single clone coverage"
     misc_feature    80993. .81045
                     /note="single clone coverage"
     misc_feature    158760. .158799
                     /note="single clone coverage"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 181792;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      35625 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 35574


RESULT 161
AC067929/c
LOCUS       AC067929               182152 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 6 clone RP11-403I20 map 6, WORKING DRAFT
            SEQUENCE, 18 unordered pieces.
ACCESSION   AC067929
VERSION     AC067929.2  GI:8247824
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 182152)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 6, clone RP11-403I20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182152)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 182152)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:7655991.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L10103
                Center clone name: 403_I_20
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 173395 bases at least Q40
                Consensus quality: 177417 bases at least Q30
                Consensus quality: 179105 bases at least Q20
                Insert size: 179000; agarose-fp
                Insert size: 180452; sum-of-contigs
                Quality coverage: 4.8 in Q20 bases; agarose-fp
                Quality coverage: 4.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2222: contig of 2222 bp in length
            *     2223     2322: gap of 100 bp
            *     2323     3785: contig of 1463 bp in length
            *     3786     3885: gap of 100 bp
            *     3886     6753: contig of 2868 bp in length
            *     6754     6853: gap of 100 bp
            *     6854    10036: contig of 3183 bp in length
            *    10037    10136: gap of 100 bp
            *    10137    14537: contig of 4401 bp in length
            *    14538    14637: gap of 100 bp
            *    14638    21177: contig of 6540 bp in length
            *    21178    21277: gap of 100 bp
            *    21278    30240: contig of 8963 bp in length
            *    30241    30340: gap of 100 bp
            *    30341    40123: contig of 9783 bp in length
            *    40124    40223: gap of 100 bp
            *    40224    47999: contig of 7776 bp in length
            *    48000    48099: gap of 100 bp
            *    48100    58235: contig of 10136 bp in length
            *    58236    58335: gap of 100 bp
            *    58336    67041: contig of 8706 bp in length
            *    67042    67141: gap of 100 bp
            *    67142    75921: contig of 8780 bp in length
            *    75922    76021: gap of 100 bp
            *    76022    86527: contig of 10506 bp in length
            *    86528    86627: gap of 100 bp
            *    86628   101334: contig of 14707 bp in length
            *   101335   101434: gap of 100 bp
            *   101435   117630: contig of 16196 bp in length
            *   117631   117730: gap of 100 bp
            *   117731   138293: contig of 20563 bp in length
            *   138294   138393: gap of 100 bp
            *   138394   157841: contig of 19448 bp in length
            *   157842   157941: gap of 100 bp
            *   157942   182152: contig of 24211 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182152
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6"
                     /clone="RP11-403I20"
```

```
                /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .2222
                /note="assembly_fragment"
     gap             2223. .2322
                /estimated_length=100
     misc_feature    2323. .3785
                /note="assembly_fragment"
     gap             3786. .3885
                /estimated_length=100
     misc_feature    3886. .6753
                /note="assembly_fragment"
     gap             6754. .6853
                /estimated_length=100
     misc_feature    6854. .10036
                /note="assembly_fragment"
     gap             10037. .10136
                /estimated_length=100
     misc_feature    10137. .14537
                /note="assembly_fragment"
     gap             14538. .14637
                /estimated_length=100
     misc_feature    14638. .21177
                /note="assembly_fragment"
     gap             21178. .21277
                /estimated_length=100
     misc_feature    21278. .30240
                /note="assembly_fragment"
     gap             30241. .30340
                /estimated_length=100
     misc_feature    30341. .40123
                /note="assembly_fragment"
     gap             40124. .40223
                /estimated_length=100
     misc_feature    40224. .47999
                /note="assembly_fragment
                clone_end:SP6
                vector_side:left"
     gap             48000. .48099
                /estimated_length=100
     misc_feature    48100. .58235
                /note="assembly_fragment"
     gap             58236. .58335
                /estimated_length=100
     misc_feature    58336. .67041
                /note="assembly_fragment"
     gap             67042. .67141
                /estimated_length=100
     misc_feature    67142. .75921
                /note="assembly_fragment"
     gap             75922. .76021
                /estimated_length=100
     misc_feature    76022. .86527
                /note="assembly_fragment"
     gap             86528. .86627
                /estimated_length=100
     misc_feature    86628. .101334
                /note="assembly_fragment"
     gap             101335. .101434
                /estimated_length=100
     misc_feature    101435. .117630

  Query Match             1.7%;  Score 52;  DB 14;  Length 182152;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14779 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 14728


RESULT 162
AC034244
LOCUS       AC034244              182892 bp    DNA     linear   PRI 20-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone RP11-101B14, complete sequence.
ACCESSION   AC034244
VERSION     AC034244.6  GI:14971194
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 182892)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182892)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 182892)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 182892)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 20, 2001 this sequence version replaced gi:14550305.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.3.
FEATURES             Location/Qualifiers
     source          1. .182892
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-101B14"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 182892;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91960 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 92011


RESULT 163
AC019059/c
LOCUS       AC019059              183556 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-125F14, WORKING DRAFT
            SEQUENCE, 31 unordered pieces.
ACCESSION   AC019059
VERSION     AC019059.4  GI:8567959
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 183556)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 183556)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
```

```
JOURNAL   Submitted (30-DEC-1999) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
COMMENT   On Jun 15, 2000 this sequence version replaced gi:7684541.

          ---------- Genome Center ----------
          Center: Washington University Genome Sequencing Center
          Center code: WUGSC
          Web site:http://genome.wustl.edu/gsc/index.shtml
          ---------- Project Information ----------
          Center project name: H_NH0125F14
          ---------- Summary Statistics ----------
          Sequencing vector: M13; 70%
          Sequencing vector: plasmid; 30%
          Chemistry: Dye-primer ET; 70% of reads
          Chemistry: Dye-terminator Big Dye; 30% of reads
          Assembly program: Phrap; version 0.990319
          Consensus quality: 164779 bases at least Q40
          Consensus quality: 170994 bases at least Q30
          Consensus quality: 174798 bases at least Q20
          Insert size: 167000; agarose-fp
          Insert size: 180556; sum-of-contigs
          Quality coverage: 3.76 in Q20 bases; agarose-fp
          Quality coverage: 3.61 in Q20 bases; sum-of-contigs
          ----------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 31 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     1181: contig of 1181 bp in length
          *     1182     1281: gap of unknown length
          *     1282     2783: contig of 1502 bp in length
          *     2784     2883: gap of unknown length
          *     2884     4618: contig of 1735 bp in length
          *     4619     4718: gap of unknown length
          *     4719     6318: contig of 1600 bp in length
          *     6319     6418: gap of unknown length
          *     6419     8952: contig of 2534 bp in length
          *     8953     9052: gap of unknown length
          *     9053    11436: contig of 2384 bp in length
          *    11437    11536: gap of unknown length
          *    11537    14108: contig of 2572 bp in length
          *    14109    14208: gap of unknown length
          *    14209    16184: contig of 1976 bp in length
          *    16185    16284: gap of unknown length
          *    16285    20536: contig of 4252 bp in length
          *    20537    20636: gap of unknown length
          *    20637    25264: contig of 4628 bp in length
          *    25265    25364: gap of unknown length
          *    25365    29797: contig of 4433 bp in length
          *    29798    29897: gap of unknown length
          *    29898    33626: contig of 3729 bp in length
          *    33627    33726: gap of unknown length
          *    33727    37341: contig of 3615 bp in length
          *    37342    37441: gap of unknown length
          *    37442    41642: contig of 4201 bp in length
          *    41643    41742: gap of unknown length
          *    41743    47291: contig of 5549 bp in length
          *    47292    47391: gap of unknown length
          *    47392    50826: contig of 3435 bp in length
          *    50827    50926: gap of unknown length
          *    50927    58180: contig of 7254 bp in length
          *    58181    58280: gap of unknown length
          *    58281    63165: contig of 4885 bp in length
          *    63166    63265: gap of unknown length
          *    63266    68481: contig of 5216 bp in length
          *    68482    68581: gap of unknown length
          *    68582    73841: contig of 5260 bp in length
          *    73842    73941: gap of unknown length
          *    73942    79919: contig of 5978 bp in length
          *    79920    80019: gap of unknown length
          *    80020    86721: contig of 6702 bp in length
          *    86722    86821: gap of unknown length
          *    86822    92876: contig of 6055 bp in length
          *    92877    92976: gap of unknown length
          *    92977    98096: contig of 5120 bp in length
          *    98097    98196: gap of unknown length
          *    98197   106663: contig of 8467 bp in length
          *   106664   106763: gap of unknown length
          *   106764   117153: contig of 10390 bp in length
          *   117154   117253: gap of unknown length
          *   117254   128332: contig of 11079 bp in length
          *   128333   128432: gap of unknown length
          *   128433   139368: contig of 10936 bp in length
          *   139369   139468: gap of unknown length
          *   139469   154140: contig of 14672 bp in length
          *   154141   154240: gap of unknown length
          *   154241   167723: contig of 13483 bp in length
          *   167724   167823: gap of unknown length
          *   167824   183556: contig of 15733 bp in length.
FEATURES             Location/Qualifiers
     source          1. .183556
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP11-125F14"
     misc_feature    1. .1181
                     /note="assembly_name:Contig7"
     gap             1182. .1281
                     /estimated_length=unknown
     misc_feature    1282. .2783
                     /note="assembly_name:Contig8"
     gap             2784. .2883
                     /estimated_length=unknown
     misc_feature    2884. .4618
                     /note="assembly_name:Contig9"
     gap             4619. .4718
                     /estimated_length=unknown
     misc_feature    4719. .6318
                     /note="assembly_name:Contig10"
     gap             6319. .6418
                     /estimated_length=unknown
     misc_feature    6419. .8952
                     /note="assembly_name:Contig11"
     gap             8953. .9052
                     /estimated_length=unknown
     misc_feature    9053. .11436
                     /note="assembly_name:Contig12"
     gap             11437. .11536
                     /estimated_length=unknown
     misc_feature    11537. .14108
                     /note="assembly_name:Contig13"
     gap             14109. .14208
                     /estimated_length=unknown
     misc_feature    14209. .16184
                     /note="assembly_name:Contig14"
     gap             16185. .16284
                     /estimated_length=unknown
     misc_feature    16285. .20536
                     /note="assembly_name:Contig15"
     gap             20537. .20636
                     /estimated_length=unknown
     misc_feature    20637. .25264
                     /note="assembly_name:Contig16"
     gap             25265. .25364
                     /estimated_length=unknown
     misc_feature    25365. .29797
                     /note="assembly_name:Contig17"
     gap             29798. .29897
                     /estimated_length=unknown
     misc_feature    29898. .33626
```

```
                     /note="assembly_name:Contig18"
-    gap             33627. .33726
                     /estimated_length=unknown
     misc_feature    33727. .37341
                     /note="assembly_name:Contig19"
     gap             37342. .37441
                     /estimated_length=unknown
     misc_feature    37442. .41642
                     /note="assembly_name:Contig20"
     gap             41643. .41742
                     /estimated_length=unknown
     misc_feature    41743. .47291
                     /note="assembly_name:Contig21"
     gap             47292. .47391
                     /estimated_length=unknown
     misc_feature    47392. .50826
                     /note="assembly_name:Contig22"
     gap             50827. .50926
                     /estimated_length=unknown
     misc_feature    50927. .58180
                     /note="assembly_name:Contig23"
     gap             58181. .58280
                     /estimated_length=unknown
     misc_feature    58281. .63165
                     /note="assembly_name:Contig24"
     gap             63166. .63265
                     /estimated_length=unknown
     misc_feature    63266. .68481
                     /note="assembly_name:Contig25"
     gap             68482. .68581
                     /estimated_length=unknown
     misc_feature    68582. .73841
                     /note="assembly_name:Contig26"
     gap             73842. .73941
                     /estimated_length=unknown
     misc_feature    73942. .79919

  Query Match             1.7%;  Score 52;  DB 14;  Length 183556;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127068 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 127017


RESULT 164
AC139565/c
LOCUS       AC139565               184252 bp    DNA     linear   HTG 20-FEB-2003
DEFINITION  Homo sapiens chromosome 15 clone RP13-620J1 map 15, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC139565
VERSION     AC139565.2  GI:28436285
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 184252)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP13-620J1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184252)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 184252)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 20, 2003 this sequence version replaced gi:28209707.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L29407
                Center clone name: 620_J_1
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 183540 bases at least Q40
                Consensus quality: 183745 bases at least Q30
                Consensus quality: 183810 bases at least Q20
                Insert size: 181000; agarose-fp
                Insert size: 183952; sum-of-contigs
                Quality coverage: 18.0 in Q20 bases; agarose-fp
                Quality coverage: 17.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    89178: contig of 89178 bp in length
            *    89179    89278: gap of 100 bp
            *    89279    89403: contig of 125 bp in length
```

```
            *    89404    89503: gap of 100 bp
            *    89504    90667: contig of 1164 bp in length
            *    90668    90767: gap of 100 bp
            *    90768   184252: contig of 93485 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184252
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP13-620J1"
                     /clone_lib="RPCI-13 Human Female BAC"
     misc_feature    1. .89178
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             89179. .89278
                     /estimated_length=100
     misc_feature    89279. .89403
                     /note="assembly_fragment"
     gap             89404. .89503
                     /estimated_length=100
     misc_feature    89504. .90667
                     /note="assembly_fragment"
     gap             90668. .90767
                     /estimated_length=100
     misc_feature    90768. .184252
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 184252;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154414 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 154363


RESULT 165
AC008675
LOCUS       AC008675              186415 bp    DNA     linear   PRI 20-DEC-2000
DEFINITION  Homo sapiens chromosome 5 clone CTB-45I8, complete sequence.
ACCESSION   AC008675
VERSION     AC008675.5  GI:11908270
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186415)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186415)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 186415)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Dec 20, 2000 this sequence version replaced gi:7709256.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 1.3.
FEATURES             Location/Qualifiers
     source          1. .186415
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-45I8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 186415;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116437 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 116488


RESULT 166
AC018757/c
LOCUS       AC018757              186418 bp    DNA     linear   PRI 03-OCT-2001
DEFINITION  Homo sapiens chromosome 5 clone CTB-45J8, complete sequence.
ACCESSION   AC018757
VERSION     AC018757.6  GI:15887297
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186418)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186418)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 186418)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 186418)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Oct 3, 2001 this sequence version replaced gi:9625331.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.4% of Sequence;
            Estimated Total Number of Errors is 1.1.
FEATURES             Location/Qualifiers
     source          1. .186418
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-45J8"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 186418;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      69981 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 69930


RESULT 167
AC034129
LOCUS       AC034129               186448 bp    DNA     linear   HTG 09-JUN-2001
DEFINITION  Homo sapiens chromosome 12 clone RP11-254M16 map 12, WORKING DRAFT
            SEQUENCE, 27 unordered pieces.
ACCESSION   AC034129
VERSION     AC034129.2  GI:7770557
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 186448)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 12, clone RP11-254M16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186448)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 12, 2000 this sequence version replaced gi:7408060.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9041
                Center clone name: 254_M_16
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 171513 bases at least Q40
                Consensus quality: 178410 bases at least Q30
                Consensus quality: 181612 bases at least Q20
                Insert size: 183848; sum-of-contigs
                Quality coverage: 4.1 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 27 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *      1     1344: contig of 1344 bp in length
            *   1345     1444: gap of 100 bp
            *   1445     2840: contig of 1396 bp in length
            *   2841     2940: gap of 100 bp
            *   2941     4409: contig of 1469 bp in length
            *   4410     4509: gap of 100 bp
            *   4510     6019: contig of 1510 bp in length
            *   6020     6119: gap of 100 bp
            *   6120     8353: contig of 2234 bp in length
            *   8354     8453: gap of 100 bp
            *   8454    11512: contig of 3059 bp in length
            *  11513    11612: gap of 100 bp
            *  11613    13961: contig of 2349 bp in length
            *  13962    14061: gap of 100 bp
            *  14062    15526: contig of 1465 bp in length
            *  15527    15626: gap of 100 bp
            *  15627    18439: contig of 2813 bp in length
            *  18440    18539: gap of 100 bp
            *  18540    20286: contig of 1747 bp in length
            *  20287    20386: gap of 100 bp
            *  20387    24011: contig of 3625 bp in length
            *  24012    24111: gap of 100 bp
            *  24112    27768: contig of 3657 bp in length
            *  27769    27868: gap of 100 bp
            *  27869    31346: contig of 3478 bp in length
            *  31347    31446: gap of 100 bp
            *  31447    35421: contig of 3975 bp in length
            *  35422    35521: gap of 100 bp
            *  35522    40318: contig of 4797 bp in length
            *  40319    40418: gap of 100 bp
            *  40419    47211: contig of 6793 bp in length
            *  47212    47311: gap of 100 bp
            *  47312    54425: contig of 7114 bp in length
            *  54426    54525: gap of 100 bp
            *  54526    61784: contig of 7259 bp in length
            *  61785    61884: gap of 100 bp
            *  61885    69081: contig of 7197 bp in length
            *  69082    69181: gap of 100 bp
            *  69182    77835: contig of 8654 bp in length
            *  77836    77935: gap of 100 bp
            *  77936    87901: contig of 9966 bp in length
            *  87902    88001: gap of 100 bp
            *  88002   100327: contig of 12326 bp in length
            * 100328   100427: gap of 100 bp
            * 100428   110556: contig of 10129 bp in length
            * 110557   110656: gap of 100 bp
            * 110657   125072: contig of 14416 bp in length
            * 125073   125172: gap of 100 bp
            * 125173   138846: contig of 13674 bp in length
            * 138847   138946: gap of 100 bp
            * 138947   162305: contig of 23359 bp in length
            * 162306   162405: gap of 100 bp
            * 162406   186448: contig of 24043 bp in length.
FEATURES             Location/Qualifiers
     source          1. .186448
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12"
                     /clone="RP11-254M16"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1344
                     /note="assembly_fragment"
     gap             1345. .1444
                     /estimated_length=100
     misc_feature    1445. .2840
                     /note="assembly_fragment"
     gap             2841. .2940
                     /estimated_length=100
     misc_feature    2941. .4409
```

```
                     /note="assembly_fragment"
     gap             4410. .4509
                     /estimated_length=100
     misc_feature    4510. .6019
                     /note="assembly_fragment"
     gap             6020. .6119
                     /estimated_length=100
     misc_feature    6120. .8353
                     /note="assembly_fragment"
     gap             8354. .8453
                     /estimated_length=100
     misc_feature    8454. .11512
                     /note="assembly_fragment"
     gap             11513. .11612
                     /estimated_length=100
     misc_feature    11613. .13961
                     /note="assembly_fragment"
     gap             13962. .14061
                     /estimated_length=100
     misc_feature    14062. .15526
                     /note="assembly_fragment"
     gap             15527. .15626
                     /estimated_length=100
     misc_feature    15627. .18439
                     /note="assembly_fragment"
     gap             18440. .18539
                     /estimated_length=100
     misc_feature    18540. .20286
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             20287. .20386
                     /estimated_length=100
     misc_feature    20387. .24011
                     /note="assembly_fragment"
     gap             24012. .24111
                     /estimated_length=100
     misc_feature    24112. .27768
                     /note="assembly_fragment"
     gap             27769. .27868
                     /estimated_length=100
     misc_feature    27869. .31346
                     /note="assembly_fragment"
     gap             31347. .31446
                     /estimated_length=100
     misc_feature    31447. .35421
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             35422. .35521
                     /estimated_length=100
     misc_feature    35522. .40318
                     /note="assembly_fragment"
     gap             40319. .40418
                     /estimated_length=100
     misc_feature    40419. .47211
                     /note="assembly_fragment"
     gap             47212. .47311
                     /estimated_length=100

  Query Match             1.7%;  Score 52;  DB 14;  Length 186448;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148872 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 148923


RESULT 168
AC120045
LOCUS       AC120045              188596 bp    DNA     linear   PRI 11-FEB-2003
DEFINITION  Homo sapiens chromosome 15, clone RP11-932O9, complete sequence.
ACCESSION   AC120045
VERSION     AC120045.19  GI:28301998
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 188596)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-932O9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 188596)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 188596)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 188596)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
```

```
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 11, 2003 this sequence version replaced gi:27733957.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L26756
                Center clone name: 932_O_9
            --------------.
FEATURES             Location/Qualifiers
     source          1. .188596
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-932O9"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(1. .362)
                     /rpt_family="L1PA15-16"
     repeat_region   complement(363. .1097)
                     /rpt_family="L1PA8"
     repeat_region   complement(1152. .1616)
                     /rpt_family="L1PA8"
     repeat_region   complement(1616. .2183)
                     /rpt_family="L1PA8"
     repeat_region   2198. .4496
                     /rpt_family="L1P"
     repeat_region   complement(4499. .5744)
                     /rpt_family="L1PA8"
     repeat_region   5899. .5921
                     /rpt_family="AT_rich"
     repeat_region   5993. .6243
                     /rpt_family="L1MC/D"
     repeat_region   6322. .6493
                     /rpt_family="MIR"
     repeat_region   complement(6832. .7032)
                     /rpt_family="MER20"
     repeat_region   7175. .7212
                     /rpt_family="(TG)n"
     repeat_region   7365. .7645
                     /rpt_family="AluSg"
     repeat_region   complement(7646. .7746)
                     /rpt_family="L1MC4a"
     repeat_region   8077. .8255
                     /rpt_family="MLT1L"
     repeat_region   complement(8615. .8709)
                     /rpt_family="MIR"
     repeat_region   complement(8718. .9169)
                     /rpt_family="LOR1b"
     repeat_region   10248. .10570
                     /rpt_family="AluSq"
     repeat_region   complement(11192. .11264)
                     /rpt_family="MIR"
     repeat_region   complement(11290. .11582)
                     /rpt_family="AluJb"
     repeat_region   complement(11606. .12358)
                     /rpt_family="L1MA3"
     repeat_region   complement(12612. .14605)
                     /rpt_family="L1MB2"
     repeat_region   complement(14608. .14911)
                     /rpt_family="AluJo"
     repeat_region   15671. .15748
                     /rpt_family="MER116"
     repeat_region   complement(16155. .16281)
                     /rpt_family="FLAM_C"
     repeat_region   16685. .17005
                     /rpt_family="AluSc"
     repeat_region   17157. .17208
                     /rpt_family="MER3"
     repeat_region   complement(17168. .17217)
                     /rpt_family="MER3"
     repeat_region   17218. .17496
                     /rpt_family="AluJb"
     repeat_region   complement(17497. .17654)
                     /rpt_family="MER3"
     repeat_region   complement(17766. .17959)
                     /rpt_family="AluJo"
     repeat_region   17981. .18226
                     /rpt_family="MSTA"
     repeat_region   19462. .19581
                     /rpt_family="MIR"
     repeat_region   20225. .20266
                     /rpt_family="MER2"
     repeat_region   complement(20260. .20516)
                     /rpt_family="MLT1F2"
     repeat_region   complement(20781. .20937)
                     /rpt_family="MLT1F2"
     repeat_region   20933. .21075
                     /rpt_family="MLT1F2"
     repeat_region   22083. .22186
                     /rpt_family="CT-rich"
     repeat_region   complement(22188. .22486)
                     /rpt_family="AluSx"
     repeat_region   22503. .22754
                     /rpt_family="L1ME2"
     repeat_region   22755. .23060
                     /rpt_family="AluSx"
     repeat_region   23063. .23297
                     /rpt_family="AluY"
     repeat_region   23298. .23351
                     /rpt_family="(TA)n"
     repeat_region   23352. .23416
                     /rpt_family="L1ME2"
     repeat_region   23449. .23580
                     /rpt_family="L1MC5"
     repeat_region   complement(23690. .23807)
                     /rpt_family="FLAM_A"
     repeat_region   complement(23834. .24139)
                     /rpt_family="AluY"

  Query Match             1.7%;  Score 52;  DB 8;  Length 188596;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92058 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 92109


RESULT 169
AC146153
LOCUS       AC146153              188640 bp    DNA     linear   HTG 01-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-37C16, *** SEQUENCING IN
            PROGRESS ***, 18 unordered pieces.
ACCESSION   AC146153
VERSION     AC146153.1  GI:33387107
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 188640)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 188640)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information --------------
            Center project name: C_PT037C16
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1483: contig of 1483 bp in length
            *     1484     1583: gap of unknown length
            *     1584     3816: contig of 2233 bp in length
            *     3817     3916: gap of unknown length
            *     3917     6890: contig of 2974 bp in length
            *     6891     6990: gap of unknown length
            *     6991     9556: contig of 2566 bp in length
            *     9557     9656: gap of unknown length
            *     9657    12497: contig of 2841 bp in length
            *    12498    12597: gap of unknown length
            *    12598    15693: contig of 3096 bp in length
            *    15694    15793: gap of unknown length
            *    15794    18490: contig of 2697 bp in length
            *    18491    18590: gap of unknown length
            *    18591    24956: contig of 6366 bp in length
            *    24957    25056: gap of unknown length
            *    25057    28576: contig of 3520 bp in length
            *    28577    28676: gap of unknown length
            *    28677    34097: contig of 5421 bp in length
            *    34098    34197: gap of unknown length
            *    34198    41105: contig of 6908 bp in length
            *    41106    41205: gap of unknown length
            *    41206    53243: contig of 12038 bp in length
            *    53244    53343: gap of unknown length
            *    53344    64385: contig of 11042 bp in length
            *    64386    64485: gap of unknown length
            *    64486    77536: contig of 13051 bp in length
            *    77537    77636: gap of unknown length
            *    77637    92854: contig of 15218 bp in length
            *    92855    92954: gap of unknown length
            *    92955   185570: contig of 92616 bp in length
            *   185571   185670: gap of unknown length
            *   185671   187117: contig of 1447 bp in length
            *   187118   187217: gap of unknown length
            *   187218   188640: contig of 1423 bp in length.
FEATURES             Location/Qualifiers
     source          1. .188640
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-37C16"
     misc_feature    1. .1483
                     /note="assembly_name:Contig13"
     gap             1484. .1583
                     /estimated_length=unknown
     misc_feature    1584. .3816
                     /note="assembly_name:Contig16"
     gap             3817. .3916
                     /estimated_length=unknown
     misc_feature    3917. .6890
                     /note="assembly_name:Contig17"
     gap             6891. .6990
                     /estimated_length=unknown
     misc_feature    6991. .9556
                     /note="assembly_name:Contig18"
     gap             9557. .9656
                     /estimated_length=unknown
     misc_feature    9657. .12497
                     /note="assembly_name:Contig19"
     gap             12498. .12597
                     /estimated_length=unknown
     misc_feature    12598. .15693
                     /note="assembly_name:Contig20"
     gap             15694. .15793
                     /estimated_length=unknown
     misc_feature    15794. .18490
                     /note="assembly_name:Contig21"
     gap             18491. .18590
                     /estimated_length=unknown
     misc_feature    18591. .24956
                     /note="assembly_name:Contig22"
     gap             24957. .25056
                     /estimated_length=unknown
     misc_feature    25057. .28576
                     /note="assembly_name:Contig23"
     gap             28577. .28676
                     /estimated_length=unknown
     misc_feature    28677. .34097
                     /note="assembly_name:Contig24"
     gap             34098. .34197
                     /estimated_length=unknown
     misc_feature    34198. .41105
                     /note="assembly_name:Contig25"
     gap             41106. .41205
                     /estimated_length=unknown
     misc_feature    41206. .53243
                     /note="assembly_name:Contig26"
     gap             53244. .53343
                     /estimated_length=unknown
     misc_feature    53344. .64385
                     /note="assembly_name:Contig27"
     gap             64386. .64485
                     /estimated_length=unknown
     misc_feature    64486. .77536
                     /note="assembly_name:Contig28"
     gap             77537. .77636
                     /estimated_length=unknown
     misc_feature    77637. .92854
                     /note="assembly_name:Contig29"
     gap             92855. .92954
                     /estimated_length=unknown
     misc_feature    92955. .185570
                     /note="assembly_name:Contig30"
     gap             185571. .185670
                     /estimated_length=unknown
     misc_feature    185671. .187117
                     /note="assembly_name:Contig6"
     gap             187118. .187217
                     /estimated_length=unknown
     misc_feature    187218. .188640
                     /note="assembly_name:Contig9"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 188640;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     172431 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 172482


RESULT 170
AC091561
LOCUS       AC091561               191037 bp    DNA     linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens chromosome 8, clone RP11-388G22, complete sequence.
ACCESSION   AC091561
VERSION     AC091561.4  GI:16041368
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191037)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-388G22
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191037)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 191037)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 11, 2001 this sequence version replaced gi:15706084.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11707
                Center clone name: 388_G_22
            --------------.
FEATURES             Location/Qualifiers
     source          1. .191037
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-388G22"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   1. .525
                     /rpt_family="L1PA8"
     repeat_region   526. .768
                     /rpt_family="L1PA8"
     repeat_region   1436. .1590
                     /rpt_family="MLT1L"
     repeat_region   complement(3292. .3332)
                     /rpt_family="MSTB"
     repeat_region   3333. .3634
                     /rpt_family="AluSx"
     repeat_region   complement(3635. .4004)
                     /rpt_family="MSTB"
     repeat_region   4057. .4201
                     /rpt_family="MIR"
     repeat_region   5412. .5507
                     /rpt_family="MIR"
     repeat_region   6261. .6392
                     /rpt_family="MER5A"
     repeat_region   6615. .6698
                     /rpt_family="L2"
     repeat_region   8232. .8255
                     /rpt_family="AT_rich"
     repeat_region   8276. .8323
                     /rpt_family="AT_rich"
     repeat_region   complement(8427. .8585)
                     /rpt_family="L2"
     repeat_region   complement(8979. .9169)
                     /rpt_family="L1ME2"
     repeat_region   complement(9224. .9495)
                     /rpt_family="L1ME"
     repeat_region   complement(9539. .10754)
                     /rpt_family="L1ME"
     repeat_region   complement(10892. .11295)
                     /rpt_family="L1ME"
     repeat_region   11324. .11633
                     /rpt_family="AluY"
     repeat_region   complement(12128. .12551)
                     /rpt_family="MLT1D"
     repeat_region   13172. .13199
                     /rpt_family="AT_rich"
     repeat_region   complement(13411. .13742)
                     /rpt_family="L2"
     repeat_region   complement(13801. .14413)
                     /rpt_family="L2"
     repeat_region   15622. .15941
                     /rpt_family="MER34B"
     repeat_region   complement(15942. .16244)
                     /rpt_family="AluJo"
     repeat_region   16245. .16465
                     /rpt_family="MER34B"
     repeat_region   17295. .17445
                     /rpt_family="MIR"
     repeat_region   17446. .17473
                     /rpt_family="AT_rich"
     repeat_region   17565. .17587
                     /rpt_family="AT_rich"
```

```
     repeat_region   complement(17591. .17704)
                     /rpt_family="FLAM_C"
     repeat_region   complement(18659. .19027)
                     /rpt_family="LTR16A"
     repeat_region   21261. .21373
                     /rpt_family="GA-rich"
     repeat_region   21496. .21526
                     /rpt_family="(TG)n"
     repeat_region   complement(21835. .22343)
                     /rpt_family="L1MEd"
     repeat_region   22937. .22971
                     /rpt_family="AT_rich"
     repeat_region   23094. .23121
                     /rpt_family="AT_rich"
     repeat_region   complement(23753. .23955)
                     /rpt_family="MIR"
     repeat_region   complement(24750. .25038)
                     /rpt_family="AluY"
     repeat_region   complement(25410. .25582)
                     /rpt_family="MIR"
     repeat_region   complement(25676. .26047)
                     /rpt_family="Charlie2a"
     repeat_region   complement(26056. .26269)
                     /rpt_family="L1MA3"
     repeat_region   complement(26271. .29438)
                     /rpt_family="L1MA3"
     repeat_region   complement(29439. .29747)
                     /rpt_family="AluY"
     repeat_region   complement(29748. .31075)
                     /rpt_family="L1MA3"
     repeat_region   31076. .31098
                     /rpt_family="(TG)n"
     repeat_region   complement(31099. .32198)
                     /rpt_family="L1MA3"
     repeat_region   32199. .32218
                     /rpt_family="(CAAAA)n"
     repeat_region   complement(32219. .33960)
                     /rpt_family="L1MA3"
     repeat_region   34076. .34102
                     /rpt_family="(TTA)n"
     repeat_region   complement(34160. .34615)
                     /rpt_family="Charlie2a"
     repeat_region   complement(34622. .34975)
                     /rpt_family="MLT1A2"
     repeat_region   34984. .35006
                     /rpt_family="AT_rich"
     repeat_region   complement(35080. .35165)
                     /rpt_family="Charlie2a"
     repeat_region   complement(35297. .35528)
                     /rpt_family="Charlie2b"
     repeat_region   complement(35577. .35856)
                     /rpt_family="Charlie2b"
     repeat_region   complement(36738. .36785)
                     /rpt_family="L2"
     repeat_region   complement(36805. .36962)
                     /rpt_family="L3"
     repeat_region   complement(36953. .37063)
                     /rpt_family="L3"
     repeat_region   37311. .37599
                     /rpt_family="L1MC2"
     repeat_region   37635. .38296

  Query Match             1.7%;  Score 52;  DB 8;  Length 191037;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3391 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 3442


RESULT 171
AC018681/c
```

```
LOCUS       AC018681              191234 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 8 clone RP11-388G22, WORKING DRAFT
            SEQUENCE, 8 unordered pieces.
ACCESSION   AC018681
VERSION     AC018681.6  GI:8954341
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191234)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191234)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 7, 2000 this sequence version replaced gi:7231039.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0388G22
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 71%
            Sequencing vector: plasmid; 29%
            Chemistry: Dye-primer ET; 71% of reads
            Chemistry: Dye-terminator Big Dye; 29% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 186413 bases at least Q40
            Consensus quality: 188204 bases at least Q30
            Consensus quality: 189106 bases at least Q20
            Insert size: 172000; agarose-fp
            Insert size: 190534; sum-of-contigs
            Quality coverage: 5.22 in Q20 bases; agarose-fp
            Quality coverage: 5.15 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1724: contig of 1724 bp in length
            *     1725     1824: gap of unknown length
            *     1825     3510: contig of 1686 bp in length
            *     3511     3610: gap of unknown length
            *     3611    11673: contig of 8063 bp in length
            *    11674    11773: gap of unknown length
            *    11774    23956: contig of 12183 bp in length
            *    23957    24056: gap of unknown length
            *    24057    40166: contig of 16110 bp in length
            *    40167    40266: gap of unknown length
            *    40267    77292: contig of 37026 bp in length
            *    77293    77392: gap of unknown length
            *    77393   125869: contig of 48477 bp in length
            *   125870   125969: gap of unknown length
            *   125970   191234: contig of 65265 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191234
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /clone="RP11-388G22"
```

```
     misc_feature    1. .1724
                     /note="assembly_name:Contig8"
     gap             1725. .1824
                     /estimated_length=unknown
     misc_feature    1825. .3510
                     /note="assembly_name:Contig9"
     gap             3511. .3610
                     /estimated_length=unknown
     misc_feature    3611. .11673
                     /note="assembly_name:Contig10"
     gap             11674. .11773
                     /estimated_length=unknown
     misc_feature    11774. .23956
                     /note="assembly_name:Contig11
                     clone_end:SP6
                     vector_side:right"
     gap             23957. .24056
                     /estimated_length=unknown
     misc_feature    24057. .40166
                     /note="assembly_name:Contig12"
     gap             40167. .40266
                     /estimated_length=unknown
     misc_feature    40267. .77292
                     /note="assembly_name:Contig13"
     gap             77293. .77392
                     /estimated_length=unknown
     misc_feature    77393. .125869
                     /note="assembly_name:Contig14"
     gap             125870. .125969
                     /estimated_length=unknown
     misc_feature    125970. .191234
                     /note="assembly_name:Contig15
                     clone_end:T7
                     vector_side:left"
ORIGIN

. Query Match             1.7%;  Score 52;  DB 14;  Length 191234;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20572 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 20521


RESULT 172
AC144780/c
LOCUS       AC144780              191540 bp    DNA     linear   PRI 26-SEP-2003
DEFINITION  Pan troglodytes BAC clone RP43-2J12 from 7, complete sequence.
ACCESSION   AC144780
VERSION     AC144780.1  GI:30841075
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 191540)
  AUTHORS   Ali,J., Bielicki,L. and Cotton,M.
  TITLE     The sequence of Pan troglodytes BAC clone RP43-2J12
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 191540)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Sequencing of Pan troglodytes
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 191540)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 191540)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     -------------- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
              Center project name: C_PT002J12
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA
            was isolated from white blood cells obtained from a male chimpanzee
            (Pan troglodytes,'Clint', Yerkes #C0471; birthdate: 6-6-80). The
            clone and detailed information can be obtained from ResGen
            (http://www.resgen.com) or Pieter de Jong and co-workers at
            http://www.bacpac.chori.org.

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES              Location/Qualifiers
     source           1. .191540
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      /chromosome="7"
                      /map="7"
                      /clone="RP43-2J12"
                      /clone_lib="RPCI-43"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 191540;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170515 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 170464


RESULT 173
AC068676
LOCUS       AC068676              191866 bp    DNA     linear   HTG 16-DEC-2000
DEFINITION  Homo sapiens clone RP11-768N8, WORKING DRAFT SEQUENCE, 14 unordered
            pieces.
ACCESSION   AC068676
VERSION     AC068676.3  GI:11875293
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 191866)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-768N8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191866)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 16, 2000 this sequence version replaced gi:10440660.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L10325
                Center clone name: 768_N_8
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 2% of reads
                Sequencing vector: Plasmid; n/a; 98% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 186020 bases at least Q40
                Consensus quality: 188219 bases at least Q30
                Consensus quality: 189180 bases at least Q20
                Insert size: 188000; agarose-fp
                Insert size: 190566; sum-of-contigs
                Quality coverage: 9.2 in Q20 bases; agarose-fp
                Quality coverage: 9.1 in Q20 ba.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      527: contig of 527 bp in length
            *      528      627: gap of 100 bp
            *      628     1287: contig of 660 bp in length
            *     1288     1387: gap of 100 bp
            *     1388     1833: contig of 446 bp in length
            *     1834     1933: gap of 100 bp
            *     1934     2723: contig of 790 bp in length
            *     2724     2823: gap of 100 bp
            *     2824     4471: contig of 1648 bp in length
            *     4472     4571: gap of 100 bp
            *     4572     8184: contig of 3613 bp in length
            *     8185     8284: gap of 100 bp
            *     8285    18881: contig of 10597 bp in length
            *    18882    18981: gap of 100 bp
            *    18982    33829: contig of 14848 bp in length
            *    33830    33929: gap of 100 bp
            *    33930    51797: contig of 17868 bp in length
            *    51798    51897: gap of 100 bp
            *    51898    67219: contig of 15322 bp in length
            *    67220    67319: gap of 100 bp
            *    67320    88231: contig of 20912 bp in length
            *    88232    88331: gap of 100 bp
            *    88332   132689: contig of 44358 bp in length
            *   132690   132789: gap of 100 bp
            *   132790   170094: contig of 37305 bp in length
            *   170095   170194: gap of 100 bp
            *   170195   191866: contig of 21672 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191866
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-768N8"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .527
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             528. .627
                     /estimated_length=100
     misc_feature    628. .1287
                     /note="assembly_fragment"
     gap             1288. .1387
                     /estimated_length=100
     misc_feature    1388. .1833
                     /note="assembly_fragment"
     gap             1834. .1933
                     /estimated_length=100
     misc_feature    1934. .2723
                     /note="assembly_fragment"
     gap             2724. .2823
                     /estimated_length=100
     misc_feature    2824. .4471
                     /note="assembly_fragment"
     gap             4472. .4571
                     /estimated_length=100
     misc_feature    4572. .8184
                     /note="assembly_fragment"
     gap             8185. .8284
                     /estimated_length=100
     misc_feature    8285. .18881
                     /note="assembly_fragment"
     gap             18882. .18981
                     /estimated_length=100
     misc_feature    18982. .33829
                     /note="assembly_fragment"
     gap             33830. .33929
                     /estimated_length=100
     misc_feature    33930. .51797
                     /note="assembly_fragment"
     gap             51798. .51897
                     /estimated_length=100
     misc_feature    51898. .67219
                     /note="assembly_fragment"
     gap             67220. .67319
                     /estimated_length=100
     misc_feature    67320. .88231
                     /note="assembly_fragment"
     gap             88232. .88331
                     /estimated_length=100
     misc_feature    88332. .132689
                     /note="assembly_fragment"
     gap             132690. .132789
                     /estimated_length=100
     misc_feature    132790. .170094
```

```
                     /note="assembly_fragment"
     gap             170095. .170194
                     /estimated_length=100
     misc_feature    170195. .191866
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 191866;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170408 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 170459


RESULT 174
AC091440/c
LOCUS       AC091440              191924 bp    DNA     linear   HTG 26-MAY-2001
DEFINITION  Homo sapiens chromosome UNK clone RP11-561D8, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC091440
VERSION     AC091440.2  GI:14209805
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191924)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191924)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On May 26, 2001 this sequence version replaced gi:13699798.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_NH0561D08
            ---------- Summary Statistics ----------
            Sequencing vector: M13; 7%
            Sequencing vector: plasmid; 93%
            Chemistry: Dye-primer ET; 7% of reads
            Chemistry: Dye-terminator Big Dye; 93% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 177613 bases at least Q40
            Consensus quality: 182658 bases at least Q30
            Consensus quality: 185159 bases at least Q20
            Insert size: 204000; agarose-fp
            Insert size: 189524; sum-of-contigs
            Quality coverage: 3.86 in Q20 bases; agarose-fp
            Quality coverage: 4.21 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2553: contig of 2553 bp in length
            *     2554     2653: gap of unknown length
            *     2654     5221: contig of 2568 bp in length
            *     5222     5321: gap of unknown length
            *     5322     8504: contig of 3183 bp in length
            *     8505     8604: gap of unknown length
            *     8605    11131: contig of 2527 bp in length
            *    11132    11231: gap of unknown length
            *    11232    14456: contig of 3225 bp in length
            *    14457    14556: gap of unknown length
            *    14557    19951: contig of 5395 bp in length
            *    19952    20051: gap of unknown length
            *    20052    24638: contig of 4587 bp in length
            *    24639    24738: gap of unknown length
            *    24739    28963: contig of 4225 bp in length
            *    28964    29063: gap of unknown length
            *    29064    35750: contig of 6687 bp in length
            *    35751    35850: gap of unknown length
            *    35851    41905: contig of 6055 bp in length
            *    41906    42005: gap of unknown length
            *    42006    48503: contig of 6498 bp in length
            *    48504    48603: gap of unknown length
            *    48604    56276: contig of 7673 bp in length
            *    56277    56376: gap of unknown length
            *    56377    63617: contig of 7241 bp in length
            *    63618    63717: gap of unknown length
            *    63718    70545: contig of 6828 bp in length
            *    70546    70645: gap of unknown length
            *    70646    79240: contig of 8595 bp in length
            *    79241    79340: gap of unknown length
            *    79341    87568: contig of 8228 bp in length
            *    87569    87668: gap of unknown length
            *    87669    98108: contig of 10440 bp in length
            *    98109    98208: gap of unknown length
            *    98209   111884: contig of 13676 bp in length
            *   111885   111984: gap of unknown length
            *   111985   125202: contig of 13218 bp in length
            *   125203   125302: gap of unknown length
            *   125303   139501: contig of 14199 bp in length
            *   139502   139601: gap of unknown length
            *   139602   159852: contig of 20251 bp in length
            *   159853   159952: gap of unknown length
            *   159953   186323: contig of 26371 bp in length
            *   186324   186423: gap of unknown length
            *   186424   188357: contig of 1934 bp in length
            *   188358   188457: gap of unknown length
            *   188458   190113: contig of 1656 bp in length
            *   190114   190213: gap of unknown length
            *   190214   191924: contig of 1711 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191924
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="UNK"
                     /clone="RP11-561D8"
     misc_feature    1. .2553
                     /note="assembly_name:Contig10"
     gap             2554. .2653
                     /estimated_length=unknown
     misc_feature    2654. .5221
                     /note="assembly_name:Contig11"
     gap             5222. .5321
                     /estimated_length=unknown
     misc_feature    5322. .8504
                     /note="assembly_name:Contig12"
     gap             8505. .8604
                     /estimated_length=unknown
     misc_feature    8605. .11131
                     /note="assembly_name:Contig13"
     gap             11132. .11231
                     /estimated_length=unknown
     misc_feature    11232. .14456
                     /note="assembly_name:Contig14"
     gap             14457. .14556
```

```
                     /estimated_length=unknown
     misc_feature    14557. .19951
                     /note="assembly_name:Contig15"
     gap             19952. .20051
                     /estimated_length=unknown
     misc_feature    20052. .24638
                     /note="assembly_name:Contig16"
     gap             24639. .24738
                     /estimated_length=unknown
     misc_feature    24739. .28963
                     /note="assembly_name:Contig17"
     gap             28964. .29063
                     /estimated_length=unknown
     misc_feature    29064. .35750
                     /note="assembly_name:Contig18"
     gap             35751. .35850
                     /estimated_length=unknown
     misc_feature    35851. .41905
                     /note="assembly_name:Contig19"
     gap             41906. .42005
                     /estimated_length=unknown
     misc_feature    42006. .48503
                     /note="assembly_name:Contig20"
     gap             48504. .48603
                     /estimated_length=unknown
     misc_feature    48604. .56276
                     /note="assembly_name:Contig21"
     gap             56277. .56376
                     /estimated_length=unknown
     misc_feature    56377. .63617
                     /note="assembly_name:Contig22"
     gap             63618. .63717
                     /estimated_length=unknown
     misc_feature    63718. .70545
                     /note="assembly_name:Contig23"
     gap             70546. .70645
                     /estimated_length=unknown
     misc_feature    70646. .79240
                     /note="assembly_name:Contig24"
     gap             79241. .79340
                     /estimated_length=unknown
     misc_feature    79341. .87568
                     /note="assembly_name:Contig25"
     gap             87569. .87668
                     /estimated_length=unknown
     misc_feature    87669. .98108
                     /note="assembly_name:Contig26"
     gap             98109. .98208
                     /estimated_length=unknown
     misc_feature    98209. .111884
                     /note="assembly_name:Contig27"
     gap             111885. .111984
                     /estimated_length=unknown
     misc_feature    111985. .125202
                     /note="assembly_name:Contig28"
     gap             125203. .125302
                     /estimated_length=unknown
     misc_feature    125303. .139501
                     /note="assembly_name:Contig29"
     gap             139502. .139601
                     /estimated_length=unknown
     misc_feature    139602. .159852
                     /note="assembly_name:Contig30"
     gap             159853. .159952
                     /estimated_length=unknown
     misc_feature    159953. .186323
                     /note="assembly_name:Contig31
                     clone_end:T7
                     vector_side:left"
     gap             186324. .186423
                     /estimated_length=unknown
     misc_feature    186424. .188357
                     /note="assembly_name:Contig7"
     gap             188358. .188457

  Query Match             1.7%;  Score 52;  DB 14;  Length 191924;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123776 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 123725


RESULT 175
AL954212
LOCUS       AL954212              192328 bp    DNA     linear   PRI 16-OCT-2003
DEFINITION  Pan troglodytes chromosome 22 clone RP43-043L10 map 22q22.11,
            complete sequence.
ACCESSION   AL954212
VERSION     AL954212.1  GI:37605795
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 192328)
  AUTHORS   .
  CONSRTM   The Chimpanzee Chromosome 22 Sequencing Consortium
  TITLE     DNA sequence of chimpanzee chromosome 22 and its evolutionary
            implications
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 192328)
  AUTHORS   Kube,M., Sudbrak,R., Borzym,K., Lehrack,S., Thiel,J., Mueller,I.,
            Klages,S., Gimmel,V., Beck,A., Ben Kahla,A., Lehrach,H., Yaspo,M.L.
            and Reinhardt,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2002) Max-Planck-Institute for Molecular
            Genetics, Department of Vertebrate Genomics, Ihnestrasse 73,
            Berlin, 14195 Germany
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
             *Chinese National Human Genome Center at Shanghai,
              Shanghai, China;
             *GBF, Dept. of Genome Analysis, Braunschweig, Germany;
             *Institute of Molecular Biotechnology, Jena, Germany;
             *KRIBB Genome Research Center, Daejeon, Korea;
             *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
             *National Institute of Genetics, Mishima, Japan;
             *National Yang Ming University Genome Research Center, Taipei,
              Taiwan;
             *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
                 Center: Max-Planck-Institute for Molecular Genetics
                 Center code: MPIMG
            -------------- Project Information
                 Center clone name: RP43-043L10
            -------------- Summary Statistics
                 Sequencing vector: pUC18; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990329
                 Consensus quality: 192151 bases at least Q40
                 Consensus quality: 192300 bases at least Q30
                 Consensus quality: 192327 bases at least Q20
                 Quality coverage: 10.18
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. --------------
            Neighboring sequence information:
            This clone is overlapped by
                 PTB-034G05
```

```
            RP43-040C07
            ------------
            SOURCE INFORMATION:
            The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 51:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/mchimp43.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR:  pBACe3.6
            ------------.
FEATURES             Location/Qualifiers
     source          1. .192328
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="22"
                     /map="22q22.11"
                     /clone="RP43-043L10"
     source          <1. .42096
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-034G05"
                     /note="overlapping clone"
     source          153484. .>192328
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-040C07"
                     /note="overlapping clone"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 192328;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15603 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 15654


RESULT 176
AL353136
LOCUS       AL353136               192505 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-133K18 on chromosome X Contains
            a pyruvate kinase muscle (PKM2) pseudogene and the gene for
            ectodysplasin A2 isoform receptor (XEDAR), complete sequence.
ACCESSION   AL353136
VERSION     AL353136.21  GI:13274342
KEYWORDS    HTG; PKM2; XEDAR.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 192505)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 12, 2001 this sequence version replaced gi:13121368.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome X, constructed by the Sanger Centre Chromosome X Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            RP11-133K18 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .192505
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-133K18"
                     /clone_lib="RPCI-11.1"
     gene            2001. .3677
                     /locus_tag="RP11-133K18.1-001"
                     /pseudo
     CDS             2001. .3677
                     /locus_tag="RP11-133K18.1-001"
                     /note="match: proteins: AAH44007 P00548 P11974 P11979
                     P11980 P52480 Q8WUW7 Q91YI8 Q92122 Q96E76"
                     /pseudo
                     /codon_start=1
                     /product="pyruvate kinase muscle (PKM2) pseudogene"
     gene            complement(join(99890. .102346,103727. .104113,
                     106886. .107050,108674. .108759,109301. .109479,
                     120187. .120283,143248. .143294))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-001"
     mRNA            complement(join(99890. .102346,103727. .104113,
                     106886. .107050,108674. .108759,109301. .109479,
                     120187. .120283,143248. .143294))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-001"
                     /product="ectodysplasin A2 receptor"
                     /note="match: ESTs: AA631436.1 AI813293.1 BM013263.1
                     BQ772070.1 BU689709.1
                     match: cDNAs: AF298812.1 BC034919.1"
     polyA_site      complement(99890)
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-001"
     polyA_signal    complement(99908. .99913)
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-001"
     gene            complement(join(103737. .104119,104292. .104348,
                     106886. .107050,108674. .108759,109301. .109479,
                     120187. .120273))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-002"
     mRNA            complement(join(103737. .104119,104292. .104348,
                     106886. .107050,108674. .108759,109301. .109479,
                     120187. .120273))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-002"
                     /product="ectodysplasin A2 receptor"
                     /note="match: cDNAs: AY152724.1"
     CDS             complement(join(103737. .104113,106886. .107050,
                     108674. .108759,109301. .109479,120187. .120273))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-001"
```

```
                     /standard_name="OTTHUMP00000023448"
                     /note="match: proteins: Q9HAV5"
                     /codon_start=1
                     /product="ectodysplasin A2 receptor"
                     /protein_id="CAH70433.1"
                     /db_xref="GI:55663359"
                     /db_xref="GOA:Q5VYY0"
                     /db_xref="InterPro:IPR001368"
                     /db_xref="UniProt/TrEMBL:Q5VYY0"
                     /translation="MDCQENEYWDQWGRCVTCQRCGPGQELSKDCGYGEGGDAYCTAC
                     PPRRYKSSWGHHRCQSCITCAVINRVQKVNCTATSNAVCGDCLPRFYRKTRIGGLQDQ
                     ECIPCTKQTPTSEVQCAFQLSLVEADTPTVPPQEATLVALVSSLLVVFTLAFLGLFFL
                     YCKQFFNRHCQRGGLLQFEADKTAKEESLFPVPPSKETSAESQVSENIFQTQPLNPIL
                     EDDCSSTSGFPTQESFTMASCTSESHSHWVHSPIECTELDLQKFSSSASYTGAETLGG
                     NTVESTGDRLELNVPFEVPSP"
     CDS             complement(join(103737. .104119,104292. .104348,
                     106886. .107050,108674. .108759,109301. .109479,
                     120187. .120273))
                     /gene="EDA2R"
                     /locus_tag="RP11-133K18.2-002"
                     /standard_name="OTTHUMP00000023449"
                     /note="match: proteins: Q8IZA6"
                     /codon_start=1
                     /product="ectodysplasin A2 receptor"
                     /protein_id="CAH70434.1"
                     /db_xref="GI:55663360"
                     /db_xref="GOA:Q5VYX9"
                     /db_xref="InterPro:IPR001368"
                     /db_xref="UniProt/TrEMBL:Q5VYX9"
                     /translation="MDCQENEYWDQWGRCVTCQRCGPGQELSKDCGYGEGGDAYCTAC
                     PPRRYKSSWGHHRCQSCITCAVINRVQKVNCTATSNAVCGDCLPRFYRKTRIGGLQDQ
                     ECIPCTKQTPTSEVQCAFQLSLVEADTPTVPPQEATLVALVSSLLVVFTLAFLGLFFL
                     YCKQFFNRHCQREKLIIFSDPVPASLNLIPEFAGGLLQFEADKTAKEESLFPVPPSKE
                     TSAESQVSENIFQTQPLNPILEDDCSSTSGFPTQESFTMASCTSESHSHWVHSPIECT
                     ELDLQKFSSSASYTGAETLGGNTVESTGDRLELNVPFEVPSP"
ORIGIN

  Query Match              1.7%;  Score 52;  DB 8;  Length 192505;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128720 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 128771


RESULT 177
AC110601/c
LOCUS       AC110601               193351 bp    DNA     linear   HTG 27-MAR-2003
DEFINITION  Homo sapiens chromosome 15 clone RP11-422F16 map 15, 4 unordered
            pieces.
ACCESSION   AC110601
VERSION     AC110601.4  GI:25103366
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 193351)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-422F16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 193351)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 193351)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Nov 19, 2002 this sequence version replaced gi:22507177.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L25406
                Center clone name: 422_F_16
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    23397: contig of 23397 bp in length
            *    23398    23497: gap of 100 bp
            *    23498    85171: contig of 61674 bp in length
            *    85172    85271: gap of 100 bp
            *    85272   140698: contig of 55427 bp in length
            *   140699   140798: gap of 100 bp
            *   140799   193351: contig of 52553 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193351
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
```

```
                     /clone="RP11-422F16"
                     /clone_lib="RPCI-11 Human Male BAC"
     gap             23398. .23497
                     /estimated_length=100
     gap             85172. .85271
                     /estimated_length=100
     gap             140699. .140798
                     /estimated_length=100
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 193351;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183110 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 183059


RESULT 178
AC141263/c
LOCUS       AC141263               193952 bp    DNA     linear   HTG 11-MAR-2003
DEFINITION  Homo sapiens chromosome 16 clone RP11-133K18, WORKING DRAFT
            SEQUENCE, 3 unordered pieces.
ACCESSION   AC141263
VERSION     AC141263.1  GI:28913043
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 193952)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 193952)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 457695
            Center clone name: RPCI-11_133K18
            -------
            Summary Statistics
            Consensus quality: 193564 bases at least Q40
            Consensus quality: 193701 bases at least Q30
            Consensus quality: 193722 bases at least Q20
            Estimated insert size: 175000; agarose-fp estimation
            Estimated insert size: 193752; sum-of-contigs estimation
            Quality coverage:  10.71 in Q20 bases; agarose-fp estimation
            Quality coverage:  9.68 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1207: contig of 1207 bp in length
            *     1208     1307: gap of unknown length
            *     1308    74872: contig of 73565 bp in length
            *    74873    74972: gap of unknown length
            *    74973   193952: contig of 118980 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193952
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="RP11-133K18"
                     /clone_lib="RPCI human BAC library 11"
     gap             1208. .1307
                     /estimated_length=unknown
     gap             74873. .74972
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 193952;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65093 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 65042


RESULT 179
AC103705
LOCUS       AC103705               194189 bp    DNA     linear   PRI 07-JAN-2003
DEFINITION  Homo sapiens chromosome 8, clone RP11-55J15, complete sequence.
ACCESSION   AC103705
VERSION     AC103705.5  GI:27531862
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 194189)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-55J15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 194189)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 194189)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
```

```
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 194189)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jan 7, 2003 this sequence version replaced gi:24462416.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L21594
                Center clone name: 55_J_15
            --------------.
FEATURES             Location/Qualifiers
     source          1. .194189
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-55J15"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   71. .416
                     /rpt_family="AluSx"
     repeat_region   complement(773. .1076)
                     /rpt_family="AluJo"
     repeat_region   1725. .1776
                     /rpt_family="MIR"
     repeat_region   2452. .2753
                     /rpt_family="AluSg"
     repeat_region   3295. .3420
                     /rpt_family="AluSq/x"
     repeat_region   3653. .3780
                     /rpt_family="MIR"
     repeat_region   4093. .4386
                     /rpt_family="AluSx"
     repeat_region   complement(4392. .4667)
                     /rpt_family="AluJo"
     repeat_region   5046. .5088
                     /rpt_family="(TTTA)n"
     repeat_region   complement(5093. .5372)
                     /rpt_family="AluSg"
     repeat_region   5425. .5694
                     /rpt_family="L2"
     repeat_region   complement(5697. .5767)
                     /rpt_family="MIR"
     repeat_region   6202. .6432
                     /rpt_family="MIR"
     repeat_region   complement(6511. .6810)
                     /rpt_family="AluY"
     repeat_region   7006. .7115
                     /rpt_family="MIR"
     repeat_region   7276. .7753
                     /rpt_family="MLT1D"
     repeat_region   complement(10040. .10142)
                     /rpt_family="MIR"
     repeat_region   complement(10343. .10667)
                     /rpt_family="AluSx"
     repeat_region   10734. .10851
                     /rpt_family="(TA)n"
     repeat_region   complement(11143. .11453)
                     /rpt_family="AluSx"
     repeat_region   11753. .11822
                     /rpt_family="MLT1C"
     repeat_region   complement(11823. .12124)
                     /rpt_family="AluSx"
     repeat_region   12125. .12516
                     /rpt_family="MLT1C"
     repeat_region   12735. .12890
                     /rpt_family="MIR"
     repeat_region   13791. .13829
                     /rpt_family="L1MA3"
     repeat_region   13830. .14147
                     /rpt_family="AluSq"
     repeat_region   14156. .14457
                     /rpt_family="AluSx"
     repeat_region   14458. .14616
                     /rpt_family="L1MA3"
     repeat_region   15897. .15952
                     /rpt_family="L2"
     repeat_region   complement(15963. .16079)
                     /rpt_family="L2"
     repeat_region   complement(16310. .16421)
                     /rpt_family="L2"
     repeat_region   16548. .16614
                     /rpt_family="MIR"
     repeat_region   16927. .16967
                     /rpt_family="(TG)n"
     repeat_region   17335. .17407
                     /rpt_family="(TCTCTG)n"
     repeat_region   17691. .17766
                     /rpt_family="MIR"
     repeat_region   18015. .18167
                     /rpt_family="MER5A"
     repeat_region   18277. .18313
                     /rpt_family="(TG)n"
     repeat_region   18425. .18468
                     /rpt_family="T-rich"
     repeat_region   18542. .18606
                     /rpt_family="CT-rich"
     repeat_region   19749. .19961
                     /rpt_family="AluJb"
     repeat_region   20552. .20630
                     /rpt_family="CT-rich"
     repeat_region   20597. .20634
                     /rpt_family="(TCTG)n"
     repeat_region   21246. .21269
                     /rpt_family="MIR"
     repeat_region   21270. .21595
                     /rpt_family="AluSx"
     repeat_region   21596. .21764
                     /rpt_family="MIR"
```

```
     repeat_region   21936. .22047
                     /rpt_family="MIR3"
     repeat_region   22184. .22268
                     /rpt_family="L2"

  Query Match              1.7%;  Score 52;  DB 8;  Length 194189;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134370 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 134421


RESULT 180
AP002026
LOCUS       AP002026               195389 bp    DNA     linear   PRI 19-MAR-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 4q22-q24, clone:429K21,
            complete sequence.
ACCESSION   AP002026
VERSION     AP002026.2  GI:29122862
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tsai,S.
  TITLE     The Complete Genomic DNA Sequence of the Human ADH Gene Complex 1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195389)
  AUTHORS   Tsai,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2000) Shih-Feng Tsai, National Yang-Ming
            University, Institute of Genetics; 155 Li-Rong St. Section 2,
            Peitou, Taipei, Taiwan 11221, Republic of China
            (E-mail:ympetsai@ym.edu.tw, URL:http://genome.ym.edu.tw/,
            Tel:886-2-28267043, Fax:886-2-28264930)
COMMENT     On Mar 18, 2003 this sequence version replaced gi:9293861.
            Quality: the expected Phred/Phrap calculated error rate (per 10kb)
            is 0.68; Estimated total number of errors is 13.39.
FEATURES             Location/Qualifiers
     source          1. .195389
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4q22-q24"
                     /clone="429K21"
ORIGIN

  Query Match              1.7%;  Score 52;  DB 8;  Length 195389;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85260 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 85311


RESULT 181
AC125238/c
LOCUS       AC125238               196622 bp    DNA     linear   PRI 30-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-711C24 from 2, complete sequence.
ACCESSION   AC125238
VERSION     AC125238.5  GI:22267884
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 196622)
  AUTHORS   Nguyen,C., Bielicki,L. and Schatzkamer,K.
  TITLE     The sequence of Homo sapiens BAC clone RP11-711C24
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 196622)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 196622)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 196622)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 196622)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 196622)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 16, 2002 this sequence version replaced gi:22138709.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0711C24
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone. This clone is
            overlapped by AC069314 and AC010138.
FEATURES             Location/Qualifiers
     source          1. .196622
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-711C24"
                     /clone_lib="RPCI-11"
     misc_feature    11079. .11345
                     /note="CpG_island (%GC=60.7, o/e=0.67, #CpGs=18)"
     misc_feature    34311. .34561
                     /note="CpG_island (%GC=57.4, o/e=1.17, #CpGs=21)"
     misc_feature    35283. .35592
                     /note="CpG_island (%GC=73.5, o/e=0.70, #CpGs=31)"
     gene            98867. .130225
                     /gene="CD28"
     mRNA            join(98867. .99140,119025. .119381,122040. .122164,
                     127176. .130225)
                     /gene="CD28"
     CDS             join(99089. .99140,119025. .119381,122040. .122164,
                     127176. .127304)
                     /gene="CD28"
                     /note="Homo sapiens CD28 antigen (Tp44) (CD28), mRNA.;
                     H_NH0711C24.1
                     This gene was based on gi(5453610)"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAY24123.1"
                     /db_xref="GI:62988736"
                     /translation="MLRLLLALNLFPSIQVTGNKILVKQSPMLVAYDNAVNLSCKYSY
                     NLFSREFRASLHKGLDSAVEVCVVYGNYSQQLQVYSKTGFNCDGKLGNESVTFYLQNL
                     YVNQTDIYFCKIEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLFPGPSKPFWVLVVVG
                     GVLACYSLLVTVAFIIFWVRSKRSRLLHSDYMNMTPRRPGPTRKHYQPYAPPRDFAAY
                     RS"
     misc_feature    171163. .171427
                     /note="CpG_island (%GC=62.3, o/e=0.80, #CpGs=20)"
     misc_feature    177943. .178329
                     /note="CpG_island (%GC=73.1, o/e=0.60, #CpGs=34)"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 196622;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70272 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 70221


RESULT 182
AC099058
LOCUS       AC099058              196623 bp    DNA     linear   PRI 26-JAN-2002
DEFINITION  Homo sapiens chromosome 3 clone RP11-714G12, complete sequence.
ACCESSION   AC099058 AC026268
VERSION     AC099058.2  GI:18376904
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 196623)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196623)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 196623)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Jan 26, 2002 this sequence version replaced gi:16799016.
            -------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
                Drafting Center: BCM
            -------------- Project Information
                Center project name: chr-3
                Center clone name: RP11-714G12 (bc0534)
            ------------- Summary Statistics
                Sequencing vector: unknown; 46% of reads
                Sequencing vector: plasmid; 54% of reads
                Chemistry: Dye-terminator ET; 92% of reads
                Chemistry: Dye-terminator Big Dye; 8% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 196528 bases at least Q40
                Consensus quality: 196607 bases at least Q30
                Consensus quality: 196623 bases at least Q20
                Insert size: 196623; sum-of-contigs
                Quality coverage: 10.6x in Q20 bases; sum-of-contigs
            --------------.
            Overlapping Sequences:
                5':  RP11-757A9 (UWGC:bc0549) AC104438
                3':  RP11-45P4 (UWGC:bc0161) AC015642
            --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.
             Base-by-base quality values are not generally visible from the
             GenBank flat file format but are available as part
             of this entry's ASN.1 file.

            This sequence was finished as follows unless otherwise noted:
            all regions were either double-stranded or sequenced with an
            alternate chemistry or covered by high quality data (i.e., Phred
            quality >= 30); an attempt was made to resolve all sequencing
            problems, such as compressions and repeats; all regions were
            covered by at least one plasmid subclone or more than one M13
            subclone; and the assembly was confirmed by restriction digest.
            --------------------
            Sequence Validation:
             This sequence has been validated by Multiple Complete Digest
             fingerprinting. Comparison of the experimentally derived digest
             fragments with sequence-predicted fragments is given below.
             The electronically-digested sequence consists of both insert and
             vector, in order to accurately represent the entire circular BAC.
             Small fragments below a variable cutoff (approximately 400-800 bp)
             are not resolved in the fingerprint and hence do not appear
             in the table. There are no significant remaining discrepancies
             between the experimental and predicted values. Uniquely ordered
             fragments are separated by dashed lines.
                   EcoRI                HindIII                BglII
            ------------------   ------------------   ------------------
            SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt   SeqDerMap  FngrPrnt
            ---------  --------   ---------  --------   ---------  --------

            ---------             ---------             ---------
                 8696      8773         873       862        6586      6552
            ---------             ---------             ---------
                    6     <800        6382      6565        2067      2057
            ---------             ---------             ---------
                 1412      1360         512      <800       11841     11792
            ---------             ---------             ---------
                 3175      3219         449      <800        2788      2800
            ---------             ---------             ---------
                 4008      4037       10461     10208        1276      1274
            ---------             ---------             ---------
```

3177	3219	1048	1014	5097	5045
11	<800	48	<800	30195	30333
13	<800	13755	13252	12139	11792
18	<800	1120	1124	2717	2800
152	<800	4425	4426	15035	14805
2523	2608	243	<800	4759	4767
1226	1156	710	<800	352	<800
1584	1667	1273	1209	2393	2384
1748	1667	2375	2359	2558	2580
2184	2196	1606	1605	6014	5959
1165	1156	2656	2848	8265	8530
939	944	14802	14596	1985	1920
187	<800	3065	3205	6151	6165
2562	2608	2314	2428	9322	9390
340	<800	1401	1419	842	857
1077	1069	893	914	3925	3969
1154	1156	2615	2740	772	857
1674	1667	2173	2185	8443	8530
1185	1156	58	<800	3279	3297
26	<800	172	<800	29	<800
3681	3677	998	1014	220	<800
237	<800	1850	1898	10592	10553
4100	4037	4785	4841	1322	1274
8739	8773	126	<800	1407	1376
1356	1360	2675	2848	828	857
1528	1532	1217	1209	1138	1121
2083	2196	1445	1419	5860	5959
6682	6726	4251	4200	153	<800
8269	8217	7428	7543	250	<800
3621	3677	3474	3582	2351	2384
5344	5303	3347	3426	867	857
6076	6046	1463	1419	6660	6731
4071	4037	839	862	4484	4513
776	<800	10565	10208	559	<800
746	794	662	<800	3026	3026
3244	3219	9	<800	10591	10553
6000	6046	2925	3071	34	<800

448	<800	728	<800	3734	3747
149	<800	4397	4426	2416	2384
3017	3219	1451	1419		
315	<800	2911	3071		
299	<800	4929	5019		
247	<800	1906	1898		
8940	8773	325	<800		
1367	1360	22385	22824		
1983	1978	2144	2185		
86	<800	1909	1898		
3759	3677	10063	10208		
7559	7610	2713	2848		
3376	3447	6186	6297		
69	<800	9507	9450		
788	794	1427	1419		

Query Match 1.7%; Score 52; DB 8; Length 196623;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127069 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 127120
```

RESULT 183
AC026160
LOCUS AC026160 198105 bp DNA linear HTG 31-MAY-2000
DEFINITION Homo sapiens chromosome 3 clone RP11-1017A6 map 3p, WORKING DRAFT SEQUENCE, 16 unordered pieces.
ACCESSION AC026160
VERSION AC026160.2 GI:8121162
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 198105)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y., Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Yu,J. and Yang,H.
TITLE Chromosome 3p genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 198105)
AUTHORS Kang,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J. and Yang,H.

```
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On May 31, 2000 this sequence version replaced gi:7271977.
            --------------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            -------------- Project Information
            Center project name:1% project
            Center clone name: RP11-1017A6
            -------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye; 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 189600 bases at least Q40
            Consensus quality: 194382 bases at least Q30
            Consensus quality: 197736 bases at least Q20
            Insert size: 190445; sum-of-contigs
            Quality coverage:   4.76x in Q20 bases;sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1194: contig of 1194 bp in length
            *     1195     1294: gap of unknown length
            *     1295     3027: contig of 1733 bp in length
            *     3028     3127: gap of unknown length
            *     3128     4527: contig of 1400 bp in length
            *     4528     4627: gap of unknown length
            *     4628     8413: contig of 3786 bp in length
            *     8414     8513: gap of unknown length
            *     8514    13191: contig of 4678 bp in length
            *    13192    13291: gap of unknown length
            *    13292    19447: contig of 6156 bp in length
            *    19448    19547: gap of unknown length
            *    19548    25715: contig of 6168 bp in length
            *    25716    25815: gap of unknown length
            *    25816    36292: contig of 10477 bp in length
            *    36293    36392: gap of unknown length
            *    36393    49087: contig of 12695 bp in length
            *    49088    49187: gap of unknown length
            *    49188    61854: contig of 12667 bp in length
            *    61855    61954: gap of unknown length
            *    61955    75324: contig of 13370 bp in length
            *    75325    75424: gap of unknown length
            *    75425    88685: contig of 13261 bp in length
            *    88686    88785: gap of unknown length
            *    88786   104410: contig of 15625 bp in length
            *   104411   104510: gap of unknown length
            *   104511   120475: contig of 15965 bp in length
            *   120476   120575: gap of unknown length
            *   120576   141214: contig of 20639 bp in length
            *   141215   141314: gap of unknown length
            *   141315   198105: contig of 56791 bp in length.
FEATURES             Location/Qualifiers
     source          1. .198105
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-1017A6"
     misc_feature    1. .1194
                     /note="assembly_name:Contig3"
     gap             1195. .1294
                     /estimated_length=unknown
     misc_feature    1295. .3027
                     /note="assembly_name:Contig4"
     gap             3028. .3127
                     /estimated_length=unknown
     misc_feature    3128. .4527
                     /note="assembly_name:Contig5"
     gap             4528. .4627
                     /estimated_length=unknown
     misc_feature    4628. .8413
                     /note="assembly_name:Contig6"
     gap             8414. .8513
                     /estimated_length=unknown
     misc_feature    8514. .13191
                     /note="assembly_name:Contig7"
     gap             13192. .13291
                     /estimated_length=unknown
     misc_feature    13292. .19447
                     /note="assembly_name:Contig8
                     clone_end:SP6
                     vector_side:right"
     gap             19448. .19547
                     /estimated_length=unknown
     misc_feature    19548. .25715
                     /note="assembly_name:Contig9"
     gap             25716. .25815
                     /estimated_length=unknown
     misc_feature    25816. .36292
                     /note="assembly_name:Contig10"
     gap             36293. .36392
                     /estimated_length=unknown
     misc_feature    36393. .49087
                     /note="assembly_name:Contig11"
     gap             49088. .49187
                     /estimated_length=unknown
     misc_feature    49188. .61854
                     /note="assembly_name:Contig12
                     clone_end:T7
                     vector_side:right"
     gap             61855. .61954
                     /estimated_length=unknown
     misc_feature    61955. .75324
                     /note="assembly_name:Contig13"
     gap             75325. .75424
                     /estimated_length=unknown
     misc_feature    75425. .88685
                     /note="assembly_name:Contig14"
     gap             88686. .88785
                     /estimated_length=unknown
     misc_feature    88786. .104410
                     /note="assembly_name:Contig15"
     gap             104411. .104510
                     /estimated_length=unknown
     misc_feature    104511. .120475
                     /note="assembly_name:Contig16"
     gap             120476. .120575
                     /estimated_length=unknown
     misc_feature    120576. .141214
                     /note="assembly_name:Contig17"
     gap             141215. .141314
                     /estimated_length=unknown
     misc_feature    141315. .198105
                     /note="assembly_name:Contig18"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 198105;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     113417 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 113468


RESULT 184
AC008108/c
LOCUS       AC008108               200000 bp    DNA     linear   HTG 23-JUL-1999
DEFINITION  Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***, 17
            unordered pieces.
ACCESSION   AC008108
VERSION     AC008108.1  GI:5578727
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 200000)
  AUTHORS   Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 200000)
  AUTHORS   Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Department of Genetics, Stanford Human
            Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1205: contig of 1205 bp in length
            *     1206    12031: gap of unknown length
            *    12032    13178: contig of 1147 bp in length
            *    13179    24004: gap of unknown length
            *    24005    25125: contig of 1121 bp in length
            *    25126    35951: gap of unknown length
            *    35952    37175: contig of 1224 bp in length
            *    37176    48001: gap of unknown length
            *    48002    49547: contig of 1546 bp in length
            *    49548    60373: gap of unknown length
            *    60374    61456: contig of 1083 bp in length
            *    61457    72281: gap of unknown length
            *    72282    73415: contig of 1134 bp in length
            *    73416    84240: gap of unknown length
            *    84241    85493: contig of 1253 bp in length
            *    85494    96318: gap of unknown length
            *    96319    97555: contig of 1237 bp in length
            *    97556   108380: gap of unknown length
            *   108381   110360: contig of 1980 bp in length
            *   110361   121185: gap of unknown length
            *   121186   122778: contig of 1593 bp in length
            *   122779   133603: gap of unknown length
            *   133604   135588: contig of 1985 bp in length
            *   135589   146413: gap of unknown length
            *   146414   148101: contig of 1688 bp in length
            *   148102   158926: gap of unknown length
            *   158927   160582: contig of 1656 bp in length
            *   160583   171407: gap of unknown length
            *   171408   173903: contig of 2496 bp in length
            *   173904   184728: gap of unknown length
            *   184729   186679: contig of 1951 bp in length
            *   186680   197504: gap of unknown length
            *   197505   200000: contig of 2496 bp in length.
FEATURES             Location/Qualifiers
     source          1. .200000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
     gap             1206. .12031
                     /estimated_length=unknown
     gap             13179. .24004
                     /estimated_length=unknown
     gap             25126. .35951
                     /estimated_length=unknown
     gap             37176. .48001
                     /estimated_length=unknown
     gap             49548. .60373
                     /estimated_length=unknown
     gap             61457. .72281
                     /estimated_length=unknown
     gap             73416. .84240
                     /estimated_length=unknown
     gap             85494. .96318
                     /estimated_length=unknown
     gap             97556. .108380
                     /estimated_length=unknown
     gap             110361. .121185
                     /estimated_length=unknown
     gap             122779. .133603
                     /estimated_length=unknown
     gap             135589. .146413
                     /estimated_length=unknown
     gap             148102. .158926
                     /estimated_length=unknown
     gap             160583. .171407
                     /estimated_length=unknown
     gap             173904. .184728
                     /estimated_length=unknown
     gap             186680. .197504
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 200000;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96734 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 96683


RESULT 185
CNS01DSQ/c
LOCUS       CNS01DSQ               200853 bp    DNA     linear   PRI 26-SEP-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-747H7 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL121839
VERSION     AL121839.3  GI:15796542
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 200853)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 200853)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Sep 27, 2001 this sequence version replaced gi:9955595.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
```

```
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : R-818K5 (AC=AL590328)
            Downstream BAC (overlapping the SP6 end) : R-248B24 ---------------
            Summary Statistics
            Assembly program: Phrap; version 2.0
            --------------
            ----------------------
            Overall quality chart :
            Range    :   bases
            0        :
            1  -  9  :       1
            10 - 19  :      10
            20 - 29  :      37
            30 - 39  :     228
            40 - 49  :    3790
            50 - 59  :   12266
            60 - 69  :   14881
            70 - 79  :   23014
            80 - 89  :   49600
            90 - 99  :   97026
            ----------------------
            Percentage of bases with a quality value >= 40 : 99 %.
FEATURES                Location/Qualifiers
     source             1. .200853
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="14"
                        /clone="R-747H7"
                        /clone_lib="RPCI-11"
     STS                74706. .74805
                        /note="matching EMBL:R33632
                        RHdb:RH53920
                        dbSTS:STS32121
                        Identified using the e-PCR software (G. Schuler)"
     STS                82261. .82445
                        /note="matching EMBL:R25802
                        RHdb:RH44322
                        RHdb:RH11302
                        dbSTS:STS1025
                        Identified using the e-PCR software (G. Schuler)"
     STS                83496. .83749
                        /note="matching EMBL:H09564
                        RHdb:RH53633
                        dbSTS:STS42877
                        Identified using the e-PCR software (G. Schuler)"
     STS                119346. .119486
                        /note="matching EMBL:G09073
                        RHdb:RH53763
                        RHdb:RH28183
                        dbSTS:STS32298
                        Identified using the e-PCR software (G. Schuler)"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 8;  Length 200853;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94881 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 94830


RESULT 186
AL162591
LOCUS       AL162591              201167 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-350G8 on chromosome 1 Contains
            the 3' end of the AQP10 gene for aquaporin 10, the ATP8B2 gene for
            ATPase, Class I, type 8B, member 2, a mitochondrial ribosomal
            protein S33 (MRPS33) pseudogene, a novel gene, the IL6R gene for
            interleukin 6 receptor, a proteasome (prosome, macropain) 26S
            subunit, non-ATPase, 8 (PSMD8) pseduogene, a novel gene
            (LOC126669), the 5' end of a novel gene (DKFZp434M202) and a CpG
            island, complete sequence.
ACCESSION   AL162591
VERSION     AL162591.16  GI:22204151
KEYWORDS    HTG; AQP10; ATP8B2; CpG island; DKFZp434M202; IL6R; LOC126669;
            MRPS33; PSMD8.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201167)
  AUTHORS   Hall,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 11, 2002 this sequence version replaced gi:16972824.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-350G8 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES                Location/Qualifiers
     source             1. .201167
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="RZPD:RPCIB753G08350"
                        /db_xref="taxon:9606"
                        /chromosome="1"
                        /clone="RP11-350G8"
                        /clone_lib="RPCI-11.2"
     misc_feature       1
                        /note="Clone_left_end: RP11-350G8"
     gene               join(complement(2407. .2520),
                        complement(AL592078.10:38498. .44621))
                        /gene="RP11-61L14.3"
                        /locus_tag="RP11-61L14.3-001"
     mRNA               join(complement(2407. .2520),
                        complement(AL592078.10:44579. .44621),
                        complement(AL592078.10:43824. .43914),
                        complement(AL592078.10:42535. .42682),
                        complement(AL592078.10:42128. .42273),
                        complement(AL592078.10:41696. .41850),
                        complement(AL592078.10:39077. .39236),
                        complement(AL592078.10:38498. .38967))
                        /gene="RP11-61L14.3"
                        /locus_tag="RP11-61L14.3-001"
                        /product="novel protein (DKFZp434M202)"
                        /note="match: cDNAs: Em:AL713777.1"
     CDS                join(complement(2407. .2433),
```

```
                complement(AL592078.10:44579. .44621),
                complement(AL592078.10:43824. .43914),
                complement(AL592078.10:42535. .42682),
                complement(AL592078.10:42128. .42273),
                complement(AL592078.10:41696. .41850),
                complement(AL592078.10:39088. .39236))
                /gene="RP11-61L14.3"
                /locus_tag="RP11-61L14.3-001"
                /standard_name="OTTHUMP00000035375"
                /codon_start=1
                /product="novel protein (DKFZp434M202)"
                /protein_id="CAH72850.1"
                /db_xref="GI:55663063"
                /db_xref="InterPro:IPR008191"
                /db_xref="UniProt/TrEMBL:Q5VZ19"
                /translation="MIQQPRAPLVLEKASGEGFGKTAAIIQLAPKAPVDLCETEKLRA
                AFFAVPLEMRGSFLVLLLRECFRDLSWLALIHSVRGEAGLLVTSIVPKTPFFWAMHVT
                EALHQNMQALFSTLAQAEEQQPYLEGSTVMRGTRCLAEYHLGDYGHAWNRCWVLDRVD
                TWAVVMFIDFGQLATIPVQSLRSLDSDDFWTIPPLTQPFMLEKDILSSYEVVHRILKG
                KITGALNSAVTAPASNLAVVPPLLPLGCLQQAAA"
gene            join(21773. .22450,24648. .24774,34530. .34835,37203. .37359,
                37824. .37943,39551. .44405)
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-001"
mRNA            join(21773. .22450,24648. .24774,34530. .34835,37203. .37359,
                37824. .37943,39551. .44405)
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-001"
                /product="novel protein (LOC126669)"
                /note="match: ESTs: Em:AA256153.1 Em:AA331688.1
                Em:AA608795.1 Em:AI143934.1 Em:AI203653.1 Em:AI218131.1
                Em:AI218256.1 Em:AI274859.1 Em:AI635642.1 Em:AI699294.1
                Em:AI797276.1 Em:AI824144.1 Em:AW014449.1 Em:BM991263.1
                Em:BX088810.1 Em:CA436491.1 Em:CA445011.1 Em:CB265391.1
                Em:CB270316.1 Em:CD300256.1 Em:CD359900.1 Em:CF552824.1
                Em:T39217.1
                match: cDNAs: Em:AK034662.1 Em:AK074067.1 Em:AK126006.1
                Em:AK131004.1"
CDS             join(21860. .22450,24648. .24774,34530. .34835,37203. .37359,
                37824. .37943,39551. .39737)
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-001"
                /standard_name="OTTHUMP00000034319"
                /note="match: proteins: Tr:Q8BSD5 Tr:Q8TEQ5"
                /codon_start=1
                /product="novel protein (LOC126669)"
                /protein_id="CAH72851.1"
                /db_xref="GI:55663064"
                /db_xref="GOA:Q5VZ18"
                /db_xref="InterPro:IPR000980"
                /db_xref="UniProt/TrEMBL:Q5VZ18"
                /translation="MQWSPTPGASACLGWASSLACSTAPTLLGRAGRGPLMAAKWFKE
                FPLNLKTVSERAKPGGGGGKLRKNSEAGGAGPGPGKGRKNSAAELGSGRAGVGPKDSR
                LSRDSLQGLIQAAAGKGRKNSRATEEEPHRGATKSSGCSTYINRLIKVDTQEKNGKSN
                YPSSSSSSSSSSSASSSPSSLGPELDKGKIIKQQETVIILEDYADPYDAKRTKGQRD
                AERVGENDGYMEPYDAQQMITEIRRRGSKDPLVKALQLLDSPCEPADGGLKSETLAKR
                RSSKDLLGKPPQLYDTPYEPAEGGPRAEGKARPPDSRLPENDERPAAEYEQPWEWKKE
                QIVRALSVQFEGAERPSFREETVRQHHRQKSWTQKILKPALSDHSEGEKVDPGLPLEK
                QPWYHGAISRAEAESRLQPCKEAGYLVRNSESGNSRYSIALKTSQGCVHIIVAQTKDN
                KYTLNQTSAVFDSIPEVVHYYSNEKLPFKGAEHMTLLYPVHSKLH"
gene            join(37842. .37943,53524. .54114)
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-002"
mRNA            join(37842. .37943,53524. .54114)
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-002"
                /product="novel protein (LOC126669)"
                /note="match: ESTs: Em:BM974630.1"
polyA_signal    44384. .44389
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-002"
polyA_site      44405
                /gene="RP11-350G8.8"
                /locus_tag="RP11-350G8.8-002"
gene            complement(join(54436. .58752,69305. .69398,73906. .73975,
                75715. .75761,87776. .87917,88730. .88896,89186. .89367,
                93280. .93403,94442. .94690,118172. .118693))
                /gene="IL6R"
                /locus_tag="RP11-350G8.6-001"
mRNA            complement(join(54436. .58752,69305. .69398,73906. .73975,
                75715. .75761,87776. .87917,88730. .88896,89186. .89367,
                93280. .93403,94442. .94690,118172. .118693))
                /gene="IL6R"
                /locus_tag="RP11-350G8.6-001"
                /product="interleukin 6 receptor"
                /note="match: ESTs: Em:AA035013.1 Em:AA193347.1
                Em:AA411559.1 Em:AI003332.1 Em:AI434387.1 Em:AI435304.1
                Em:AI971631.1 Em:AL120677.1 Em:AW177946.1 Em:AW178188.1
                Em:AW468643.1 Em:AW503691.1 Em:BG032730.1 Em:BG500539.1
                Em:BG575786.1 Em:BI018205.1 Em:BI760931.1 Em:BI860794.1
                Em:BM015654.1 Em:BM044442.1 Em:BM128112.1 Em:BM549585.1
                Em:BM911928.1 Em:BQ223533.1 Em:BQ335482.1 Em:BQ776559.1
                Em:BQ892769.1 Em:BQ937176.1 Em:BU630287.1 Em:BU632645.1
                Em:BU848859.1 Em:BX113990.1 Em:BX509093.1 Em:BX646979.1
                Em:CA308138.1 Em:CA422880.1 Em:CA427041.1 Em:CB163591.1
                Em:CB962465.1 Em:CD516644.1 Em:CD609023.1 Em:CD609024.1
                Em:CD695637.1 Em:CF123018.1 Em:H59001.1
                match: cDNAs: Em:AF147881.1 Em:BC035327.1 Em:X12830.1
                Em:X58298.1"
polyA_site      complement(54436)
                /gene="IL6R"
                /locus_tag="RP11-350G8.6-001"
polyA_signal    complement(54451. .54456)
                /gene="IL6R"
                /locus_tag="RP11-350G8.6-001"
CDS             complement(join(58506. .58752,69305. .69398,73906. .73975,
                75715. .75761,87776. .87917,88730. .88896,89186. .89367,
                93280. .93403,94442. .94690,118172. .118256))
                /gene="IL6R"
                /locus_tag="RP11-350G8.6-001"
                /standard_name="OTTHUMP00000034320"
                /note="match: proteins: Sw:P08887"
                /codon_start=1
                /product="interleukin 6 receptor"
                /protein_id="CAH72853.1"
                /db_xref="GI:55663065"
                /db_xref="InterPro:IPR002996"
                /db_xref="InterPro:IPR003530"
                /db_xref="InterPro:IPR003961"
                /db_xref="InterPro:IPR007110"
                /db_xref="InterPro:IPR008957"

  Query Match             1.7%;  Score 52;  DB 8;  Length 201167;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122716 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 122767


RESULT 187
AC090559
LOCUS       AC090559              202539 bp    DNA     linear   PRI 29-JUL-2002
DEFINITION  Homo sapiens chromosome 11, clone RP11-750H9, complete sequence.
ACCESSION   AC090559
VERSION     AC090559.7  GI:22002211
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 202539)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-750H9
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202539)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 202539)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 202539)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 29, 2002 this sequence version replaced gi:21591871.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11075
                Center clone name: 750_H_9
            --------------.
FEATURES             Location/Qualifiers
     source          1. .202539
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-750H9"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   996. .1158
                     /rpt_family="CT-rich"
     repeat_region   2010. .2190
                     /rpt_family="(CCCCAG)n"
     repeat_region   3759. .3937
                     /rpt_family="(TCCA)n"
     repeat_region   3981. .4135
                     /rpt_family="(TCCA)n"
     repeat_region   complement(4149. .4503)
                     /rpt_family="L1MA4A"
     repeat_region   4508. .4785
                     /rpt_family="(TCCA)n"
     repeat_region   4805. .4903
                     /rpt_family="(TCCA)n"
     repeat_region   complement(4911. .4987)
                     /rpt_family="L2"
     repeat_region   5029. .5059
                     /rpt_family="(TTG)n"
     repeat_region   5521. .5599
                     /rpt_family="(TTCA)n"
     repeat_region   complement(5601. .5959)
                     /rpt_family="L1PB3"
     repeat_region   5960. .5987
                     /rpt_family="(TA)n"
     repeat_region   complement(5988. .6163)
                     /rpt_family="L1PB3"
     repeat_region   complement(6164. .6470)
                     /rpt_family="AluSq"
     repeat_region   complement(6473. .6771)
                     /rpt_family="AluJb"
     repeat_region   complement(6772. .6832)
                     /rpt_family="L1PB3"
     repeat_region   complement(7731. .7814)
                     /rpt_family="MIR"
     repeat_region   7841. .8088
                     /rpt_family="GA-rich"
     repeat_region   complement(8333. .8516)
                     /rpt_family="L1ME4A"
     repeat_region   8517. .8827
                     /rpt_family="AluY"
     repeat_region   complement(8828. .8939)
                     /rpt_family="L1ME4A"
     repeat_region   complement(8940. .9144)
                     /rpt_family="L1MA3"
     repeat_region   complement(9145. .9200)
                     /rpt_family="L1ME4A"
     repeat_region   complement(9194. .9273)
                     /rpt_family="L1ME3A"
     repeat_region   9274. .9580
                     /rpt_family="AluY"
     repeat_region   complement(9581. .9641)
                     /rpt_family="L1ME3A"
     repeat_region   9642. .9955
                     /rpt_family="AluY"
```

```
     unsure          9915. .9924
                     /note="<30 qual SNGL region"
     repeat_region   complement(9956. .10051)
                     /rpt_family="L1ME3A"
     repeat_region   complement(10052. .10356)
                     /rpt_family="AluSg"
     repeat_region   complement(10533. .10835)
                     /rpt_family="AluSx"
     repeat_region   complement(10878. .10997)
                     /rpt_family="AluJb"
     repeat_region   complement(10998. .11149)
                     /rpt_family="L1ME3A"
     repeat_region   complement(11150. .11184)
                     /rpt_family="AluSx"
     repeat_region   11185. .11208
                     /rpt_family="(TTTTA)n"
     repeat_region   complement(11209. .11495)
                     /rpt_family="AluSx"
     repeat_region   complement(11500. .11622)
                     /rpt_family="FLAM_C"
     repeat_region   11667. .11805
                     /rpt_family="AluJb"
     repeat_region   11806. .11830
                     /rpt_family="(TGG)n"
     repeat_region   11831. .11976
                     /rpt_family="AluJb"
     repeat_region   12021. .12326
                     /rpt_family="AluSq"
     repeat_region   complement(12952. .13398)
                     /rpt_family="L2"
     repeat_region   complement(13452. .13561)
                     /rpt_family="MIR"
     repeat_region   complement(13636. .13765)
                     /rpt_family="MIR"
     repeat_region   14462. .14661
                     /rpt_family="MIR"
     repeat_region   14692. .14870
                     /rpt_family="MIR"
     repeat_region   complement(15253. .15394)
                     /rpt_family="MIR"
     repeat_region   complement(15871. .15976)

  Query Match             1.7%;  Score 52;  DB 8;  Length 202539;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63712 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 63763


RESULT 188
AC151042
LOCUS       AC151042               202971 bp    DNA     linear   HTG 24-NOV-2004
DEFINITION  Callithrix jacchus clone CH259-63D17, WORKING DRAFT SEQUENCE, 15
            ordered pieces.
ACCESSION   AC151042
VERSION     AC151042.2  GI:56044586
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Callithrix jacchus (white-tufted-ear marmoset)
  ORGANISM  Callithrix jacchus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Callitrichidae; Callithrix.
REFERENCE   1  (bases 1 to 202971)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202971)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 202971)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 24, 2004 this sequence version replaced gi:51372024.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: hja
                 Center clone name: 063D17

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 199626 bases at least Q40
                 Consensus quality: 200659 bases at least Q30
                 Consensus quality: 201280 bases at least Q20
                 Insert size: 228000; agarose-fp
                 Insert size: 201571; sum-of-contigs
                 Quality coverage: 7.60x in Q20 bases; agarose-fp
                 Quality coverage: 8.60x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    19090: contig of 19090 bp in length
            *    19091    19190: gap of unknown length
            *    19191    32965: contig of 13775 bp in length
            *    32966    33065: gap of unknown length
            *    33066    37634: contig of 4569 bp in length
            *    37635    37734: gap of unknown length
            *    37735    39153: contig of 1419 bp in length
            *    39154    39253: gap of unknown length
            *    39254    46148: contig of 6895 bp in length
            *    46149    46248: gap of unknown length
            *    46249    75331: contig of 29083 bp in length
            *    75332    75431: gap of unknown length
            *    75432    79016: contig of 3585 bp in length
            *    79017    79116: gap of unknown length
            *    79117    82235: contig of 3119 bp in length
```

```
            *    82236    82335: gap of unknown length
            *    82336    87728: contig of 5393 bp in length
            *    87729    87828: gap of unknown length
            *    87829   100513: contig of 12685 bp in length
            *   100514   100613: gap of unknown length
            *   100614   123670: contig of 23057 bp in length
            *   123671   123770: gap of unknown length
            *   123771   137796: contig of 14026 bp in length
            *   137797   137896: gap of unknown length
            *   137897   161136: contig of 23240 bp in length
            *   161137   161236: gap of unknown length
            *   161237   176590: contig of 15354 bp in length
            *   176591   176690: gap of unknown length
            *   176691   202971: contig of 26281 bp in length.
FEATURES             Location/Qualifiers
     source          1. .202971
                     /organism="Callithrix jacchus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9483"
                     /clone="CH259-63D17"
                     /clone_lib="CH259"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .39045
                     /note="clone overlaps with GenBank Accession Number
                     AC151015 clone CH259-164B19 (center project name hiz)"
     misc_feature    1. .19090
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     gap             19091. .19190
                     /estimated_length=unknown
     misc_feature    19191. .32965
                     /note="assembly_fragment"
     gap             32966. .33065
                     /estimated_length=unknown
     misc_feature    33066. .37634
                     /note="assembly_fragment"
     gap             37635. .37734
                     /estimated_length=unknown
     misc_feature    37735. .39153
                     /note="assembly_fragment"
     gap             39154. .39253
                     /estimated_length=unknown
     misc_feature    39254. .46148
                     /note="assembly_fragment"
     gap             46149. .46248
                     /estimated_length=unknown
     misc_feature    46249. .75331
                     /note="assembly_fragment"
     gap             75332. .75431
                     /estimated_length=unknown
     misc_feature    75432. .79016
                     /note="assembly_fragment"
     gap             79017. .79116
                     /estimated_length=unknown
     misc_feature    79117. .82235
                     /note="assembly_fragment"
     gap             82236. .82335
                     /estimated_length=unknown
     misc_feature    82336. .87728
                     /note="assembly_fragment"
     gap             87729. .87828
                     /estimated_length=unknown
     misc_feature    87829. .100513
                     /note="assembly_fragment"
     misc_feature    98925. .202971
                     /note="clone overlaps with GenBank Accession Number
                     AC150824 clone CH259-392G13 (center project name hjb)"
     gap             100514. .100613
                     /estimated_length=unknown
     misc_feature    100614. .123670
                     /note="assembly_fragment"
     gap             123671. .123770
                     /estimated_length=unknown
     misc_feature    123771. .137796
                     /note="assembly_fragment"
     gap             137797. .137896
                     /estimated_length=unknown
     misc_feature    137897. .161136
                     /note="assembly_fragment"
     gap             161137. .161236
                     /estimated_length=unknown
     misc_feature    161237. .176590
                     /note="assembly_fragment"
     gap             176591. .176690
                     /estimated_length=unknown
     misc_feature    176691. .202971
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 202971;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     103406 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 103457


RESULT 189
AC008040/c
LOCUS       AC008040              204917 bp    DNA     linear   PRI 28-SEP-2002
DEFINITION  Homo sapiens 3 BAC RP11-379K17 (Roswell Park Cancer Institue Human
            BAC Library) complete sequence.
ACCESSION   AC008040
VERSION     AC008040.7  GI:5922025
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204917)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Garcia,D.K., Gorrell,J.H.,
            Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
            Hodgson,A., Hogues,M., Holloway,C., Hosak,H., Jackson,L.E.,
            Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondejewski,N., Kong,Y.,
            Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,
            Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G.,
            Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S., Nelson,A.,
            Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,
            Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,
            Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Simon,M.,
            Sparks,A., Stamps,A., Sucgang,R., Tabor,P., Taylor,T., Vasquez,L.,
            Vinson,R., Vo,Q., Wahbah,M., Watlington,S., Weinstock,G.,
            Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,
            Yu,W., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1999) Human Genome Sequencing Center, Department
```

```
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   7  (bases 1 to 204917)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 24, 1999 this sequence version replaced gi:5739555.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

            QUALSTAT-REPORT.
FEATURES             Location/Qualifiers
     source          1. .204917
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-379K17"
     misc_feature    complement(1. .1995)
                     /note="overlaps bases 1. .1995 of clone AC023891"
                     /function="clone overlap"
     misc_feature    766. .1132
                     /note="Region: od11d01.s1 Homo sapiens cDNA, AA810014"
     misc_feature    2456. .2779
                     /note="Region: zs55f12.s1 Homo sapiens cDNA, AA286777"
     STS             2554. .2768
                     /standard_name="G37704"
                     /db_xref="dbSTS:55264"
     repeat_region   6897. .7126
                     /rpt_family="MER20"
     misc_feature    7488. .8243
                     /note="Region: Unigene cluster containing AA400519 and
                     AA165136"
     repeat_region   11171. .11471
                     /rpt_family="AluSx"
     repeat_region   complement(13526. .13835)
                     /rpt_family="AluSq"
     misc_feature    13556. .15048
                     /note="Region: Unigene cluster containing AI683116,
                     AL048825, AA278956, and W38986"
     repeat_region   complement(14286. .14586)
                     /rpt_family="AluSx"
     misc_feature    15264. .15890
                     /note="Region: Unigene cluster containing AI042019 and
                     AI890489"
     repeat_region   15800. .15834
                     /rpt_family="AT_rich"
     repeat_region   16816. .16837
                     /rpt_family="AT_rich"
     repeat_region   16911. .17154
                     /rpt_family="MIR"
     repeat_region   complement(17196. .17318)
                     /rpt_family="MER112"
     repeat_region   17598. .17782
                     /rpt_family="MER20"
     misc_feature    18072. .19357
                     /note="Region: Unigene cluster containing AI767291,
                     AI522162, W31729 and AA354506"
     STS             18120. .18226
                     /standard_name="G24029"
                     /db_xref="dbSTS:37867"
     repeat_region   20464. .20496
                     /rpt_family="(CAAAAA)n"
     repeat_region   20576. .20882
                     /rpt_family="AluJo"
     repeat_region   complement(20890. .21061)
                     /rpt_family="MER46A"
     repeat_region   21460. .21885
                     /rpt_family="L1PA13"
     repeat_region   22470. .22760
                     /rpt_family="AluJb"
     repeat_region   complement(24017. .24179)
                     /rpt_family="Charlie1b"
     repeat_region   complement(24180. .24635)
                     /rpt_family="LTR22A"
     repeat_region   complement(24639. .24744)
                     /rpt_family="HERVK22"
     repeat_region   complement(24750. .25603)
                     /rpt_family="HERVK22"
     repeat_region   complement(25606. .26058)
                     /rpt_family="LTR22A"
     repeat_region   complement(26061. .26174)
                     /rpt_family="Charlie1b"
     repeat_region   complement(26391. .26559)
                     /rpt_family="MIR"
     repeat_region   26635. .26943
                     /rpt_family="AluSx"
     repeat_region   27836. .28013
                     /rpt_family="MLT1I"
     repeat_region   complement(28215. .28526)
                     /rpt_family="AluSq"
     repeat_region   28548. .28687
```

```
                     /rpt_family="L1M4"
     repeat_region   28688. .28981
                     /rpt_family="AluJb"
     repeat_region   28982. .29281
                     /rpt_family="L1M4"
     repeat_region   29282. .29588
                     /rpt_family="AluSx"
     repeat_region   29589. .30042
                     /rpt_family="L1M4"
     repeat_region   30055. .30213
                     /rpt_family="L1M4"

  Query Match             1.7%;  Score 52;  DB 8;  Length 204917;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     160446 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 160395


RESULT 190
AC073916/c
LOCUS       AC073916               205283 bp    DNA     linear   PRI 27-MAR-2003
DEFINITION  Homo sapiens 12 BAC RP11-408I18 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC073916
VERSION     AC073916.41  GI:29293998
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 205283)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   6  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   7  (bases 1 to 205283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Mar 27, 2003 this sequence version replaced gi:22038291.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.
```

```
          QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
          standards - estimated error rate less than 1 per 10,000 bases.
          Reports of lowest quality individual bases and measures of base
          quality are listed below. Description of the metrics can be found
          at URL:
          http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
          ml.
FEATURES             Location/Qualifiers
     source          1. .205283
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-408I18"
     repeat_region   complement(1. .116)
                     /rpt_family="MLT1J2"
     repeat_region   complement(185. .545)
                     /rpt_family="MLT1A0"
     STS             482. .856
                     /standard_name="D12S820"
     repeat_region   548. .569
                     /rpt_family="AT_rich"
     repeat_region   710. .845
                     /rpt_family="(TCCA)n"
     repeat_region   complement(846. .980)
                     /rpt_family="L2"
     repeat_region   1320. .1681
                     /rpt_family="L2"
     repeat_region   2124. .2158
                     /rpt_family="AT_rich"
     repeat_region   2327. .2353
                     /rpt_family="(T)n"
     repeat_region   complement(2835. .2914)
                     /rpt_family="MER58B"
     repeat_region   3015. .3377
                     /rpt_family="L1MC4"
     repeat_region   complement(3384. .3882)
                     /rpt_family="LTR40a"
     repeat_region   3908. .3944
                     /rpt_family="U2"
     repeat_region   3983. .4182
                     /rpt_family="L1MC4"
     repeat_region   4217. .4256
                     /rpt_family="(GA)n"
     repeat_region   5472. .5647
                     /rpt_family="(TG)n"
     repeat_region   5698. .5870
                     /rpt_family="(TG)n"
     repeat_region   complement(6009. .6188)
                     /rpt_family="MIR"
     repeat_region   6410. .6484
                     /rpt_family="MIRb"
     repeat_region   complement(6501. .6815)
                     /rpt_family="AluSx"
     repeat_region   7401. .7603
                     /rpt_family="L2"
     repeat_region   8096. .8186
                     /rpt_family="MER96B"
     repeat_region   complement(8187. .8485)
                     /rpt_family="AluSx"
     repeat_region   8486. .8515
                     /rpt_family="MER96B"
     repeat_region   9052. .9174
                     /rpt_family="L2"
     repeat_region   10287. .10476
                     /rpt_family="(CGGGG)n"
     repeat_region   10415. .10551
                     /rpt_family="G-rich"
     repeat_region   10602. .10631
                     /rpt_family="GC_rich"
     repeat_region   10658. .10684
                     /rpt_family="GC_rich"
     repeat_region   10781. .10801
                     /rpt_family="(CCCCAG)n"
     repeat_region   10835. .10859
                     /rpt_family="GC_rich"
     repeat_region   11332. .11465

  Query Match             1.7%;  Score 52;  DB 8;  Length 205283;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82961 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 82910


RESULT 191
AC023560
LOCUS       AC023560              205736 bp    DNA     linear   HTG 12-MAR-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-85I20 map 17, WORKING DRAFT
            SEQUENCE, 34 unordered pieces.
ACCESSION   AC023560
VERSION     AC023560.2  GI:7229913
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 205736)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-85I20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 205736)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 12, 2000 this sequence version replaced gi:6978256.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6741
                Center clone name: 85_I_20
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 178666 bases at least Q40
```

```
    Consensus quality: 190940 bases at least Q30
    Consensus quality: 197088 bases at least Q20
    Insert size: 194000; agarose-fp
    Insert size: 202436; sum-of-contigs
    Quality coverage: 4.2 in Q20 bases; agarose-fp
    Quality coverage: 4.0 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     1267: contig of 1267 bp in length
*    1268     1367: gap of 100 bp
*    1368     2402: contig of 1035 bp in length
*    2403     2502: gap of 100 bp
*    2503     3524: contig of 1022 bp in length
*    3525     3624: gap of 100 bp
*    3625     4913: contig of 1289 bp in length
*    4914     5013: gap of 100 bp
*    5014     6289: contig of 1276 bp in length
*    6290     6389: gap of 100 bp
*    6390     7655: contig of 1266 bp in length
*    7656     7755: gap of 100 bp
*    7756     9267: contig of 1512 bp in length
*    9268     9367: gap of 100 bp
*    9368    11138: contig of 1771 bp in length
*   11139    11238: gap of 100 bp
*   11239    13822: contig of 2584 bp in length
*   13823    13922: gap of 100 bp
*   13923    16968: contig of 3046 bp in length
*   16969    17068: gap of 100 bp
*   17069    20185: contig of 3117 bp in length
*   20186    20285: gap of 100 bp
*   20286    22634: contig of 2349 bp in length
*   22635    22734: gap of 100 bp
*   22735    27012: contig of 4278 bp in length
*   27013    27112: gap of 100 bp
*   27113    29342: contig of 2230 bp in length
*   29343    29442: gap of 100 bp
*   29443    33381: contig of 3939 bp in length
*   33382    33481: gap of 100 bp
*   33482    36562: contig of 3081 bp in length
*   36563    36662: gap of 100 bp
*   36663    40982: contig of 4320 bp in length
*   40983    41082: gap of 100 bp
*   41083    45520: contig of 4438 bp in length
*   45521    45620: gap of 100 bp
*   45621    50623: contig of 5003 bp in length
*   50624    50723: gap of 100 bp
*   50724    55610: contig of 4887 bp in length
*   55611    55710: gap of 100 bp
*   55711    61038: contig of 5328 bp in length
*   61039    61138: gap of 100 bp
*   61139    66159: contig of 5021 bp in length
*   66160    66259: gap of 100 bp
*   66260    73479: contig of 7220 bp in length
*   73480    73579: gap of 100 bp
*   73580    80554: contig of 6975 bp in length
*   80555    80654: gap of 100 bp
*   80655    89313: contig of 8659 bp in length
*   89314    89413: gap of 100 bp
*   89414    98132: contig of 8719 bp in length
*   98133    98232: gap of 100 bp
*   98233   107469: contig of 9237 bp in length
*  107470   107569: gap of 100 bp
*  107570   117873: contig of 10304 bp in length
*  117874   117973: gap of 100 bp
*  117974   130973: contig of 13000 bp in length
*  130974   131073: gap of 100 bp
*  131074   143236: contig of 12163 bp in length
*  143237   143336: gap of 100 bp
*  143337   154478: contig of 11142 bp in length
*  154479   154578: gap of 100 bp
*  154579   169967: contig of 15389 bp in length
*  169968   170067: gap of 100 bp
*  170068   187347: contig of 17280 bp in length
*  187348   187447: gap of 100 bp
*  187448   205736: contig of 18289 bp in length.
FEATURES             Location/Qualifiers
     source          1. .205736
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-85I20"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1267
                     /note="assembly_fragment"
     gap             1268. .1367
                     /estimated_length=100
     misc_feature    1368. .2402
                     /note="assembly_fragment"
     gap             2403. .2502
                     /estimated_length=100
     misc_feature    2503. .3524
                     /note="assembly_fragment"
     gap             3525. .3624
                     /estimated_length=100
     misc_feature    3625. .4913
                     /note="assembly_fragment"
     gap             4914. .5013
                     /estimated_length=100
     misc_feature    5014. .6289
                     /note="assembly_fragment"
     gap             6290. .6389
                     /estimated_length=100
     misc_feature    6390. .7655
                     /note="assembly_fragment"
     gap             7656. .7755
                     /estimated_length=100
     misc_feature    7756. .9267
                     /note="assembly_fragment"
     gap             9268. .9367
                     /estimated_length=100
     misc_feature    9368. .11138
                     /note="assembly_fragment"
     gap             11139. .11238
                     /estimated_length=100
     misc_feature    11239. .13822
                     /note="assembly_fragment"
     gap             13823. .13922
                     /estimated_length=100
     misc_feature    13923. .16968
                     /note="assembly_fragment"
     gap             16969. .17068
                     /estimated_length=100
     misc_feature    17069. .20185
                     /note="assembly_fragment"
     gap             20186. .20285
                     /estimated_length=100
     misc_feature    20286. .22634
                     /note="assembly_fragment"
     gap             22635. .22734
                     /estimated_length=100
     misc_feature    22735. .27012
                     /note="assembly_fragment"
     gap             27013. .27112
                     /estimated_length=100

 Query Match             1.7%;  Score 52;  DB 14;  Length 205736;
 Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
```

```
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115864 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 115915


RESULT 192
AC015884/c
LOCUS       AC015884               207661 bp    DNA     linear   PRI 27-OCT-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-216P6, complete sequence.
ACCESSION   AC015884
VERSION     AC015884.15  GI:24415419
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 207661)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-216P6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 207661)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 207661)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 207661)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 27, 2002 this sequence version replaced gi:22297427.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L731
                Center clone name: 216_P_6
            --------------.
FEATURES             Location/Qualifiers
     source          1. .207661
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-216P6"
                     /clone_lib="RPCI-11 Human Male BAC"
     unsure          721
                     /note="probably C"
     repeat_region   2666. .2923
                     /rpt_family="AluSx"
     repeat_region   3206. .3412
                     /rpt_family="AluJo"
     repeat_region   3591. .3888
                     /rpt_family="AluSx"
     repeat_region   complement(4652. .4964)
                     /rpt_family="AluSx"
     repeat_region   5067. .5109
                     /rpt_family="Alu"
     unsure          5106. .5112
                     /note="<30 qual SNGL region"
     repeat_region   5110. .5424
                     /rpt_family="AluY"
     repeat_region   5487. .5582
                     /rpt_family="MIR3"
     repeat_region   complement(5846. .5901)
                     /rpt_family="MIR"
     repeat_region   complement(5902. .6209)
                     /rpt_family="AluY"
     repeat_region   complement(6210. .6373)
                     /rpt_family="MIR"
     repeat_region   6847. .7140
                     /rpt_family="AluSg"
     repeat_region   7860. .7997
                     /rpt_family="AluJo"
     repeat_region   7998. .8278
                     /rpt_family="AluSg"
     repeat_region   8279. .8435
                     /rpt_family="AluJo"
     repeat_region   8436. .8467
                     /rpt_family="(CA)n"
     repeat_region   9011. .9302
                     /rpt_family="AluSq"
     repeat_region   9306. .9597
                     /rpt_family="AluSx"
     repeat_region   complement(9849. .10138)
```

```
                     /rpt_family="AluJb"
     repeat_region   complement(10139. .10273)
                     /rpt_family="AluJo"
     repeat_region   complement(10285. .10579)
                     /rpt_family="AluY"
     unsure          complement(10584. .10591)
                     /note="<30 qual SNGL region"
     repeat_region   10608. .10905
                     /rpt_family="AluSx"
     repeat_region   11349. .11651
                     /rpt_family="AluSx"
     repeat_region   11882. .12022
                     /rpt_family="AluJb"
     repeat_region   12040. .12072
                     /rpt_family="AT_rich"
     repeat_region   12711. .12736
                     /rpt_family="(CA)n"
     repeat_region   12817. .13121
                     /rpt_family="AluSx"
     repeat_region   13292. .13590
                     /rpt_family="AluY"
     repeat_region   14769. .15274
                     /rpt_family="L3"
     repeat_region   15275. .15322
                     /rpt_family="AluYb8"
     repeat_region   15335. .15636
                     /rpt_family="AluSq"
     repeat_region   complement(15692. .15846)
                     /rpt_family="MER30"
     repeat_region   complement(15847. .16142)
                     /rpt_family="AluSp"
     repeat_region   complement(16143. .16218)
                     /rpt_family="MER30"
     repeat_region   16220. .16432
                     /rpt_family="L3"
     repeat_region   16559. .16962
                     /rpt_family="MLT1K"
     repeat_region   17094. .17207
                     /rpt_family="L1MC5"
     repeat_region   17208. .17482
                     /rpt_family="AluY"
     repeat_region   17483. .17545
                     /rpt_family="L1MC5"
     repeat_region   17906. .18082
                     /rpt_family="AluJb"
     repeat_region   18152. .18462
                     /rpt_family="AluSx"
     repeat_region   complement(18647. .18889)
                     /rpt_family="MIR"
     repeat_region   complement(20440. .20741)
                     /rpt_family="AluSp"
     repeat_region   20925. .21221
                     /rpt_family="AluY"
     repeat_region   21382. .21514
                     /rpt_family="AluSx"
     repeat_region   21515. .21822
                     /rpt_family="AluSp"
     repeat_region   21823. .21998
                     /rpt_family="AluSx"
     repeat_region   22246. .22543
                     /rpt_family="AluY"
     repeat_region   complement(23168. .23474)
                     /rpt_family="AluSx"

  Query Match             1.7%;  Score 52;  DB 8;  Length 207661;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116276 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 116225
```

```
RESULT 193
AC027205/c
LOCUS       AC027205               211812 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 1 clone RP11-130F6 map 1, WORKING DRAFT
            SEQUENCE, 51 unordered pieces.
ACCESSION   AC027205
VERSION     AC027205.2  GI:7652003
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 211812)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 1, clone RP11-130F6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 211812)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 211812)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 27, 2000 this sequence version replaced gi:7331532.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
```

```
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L7762
    Center clone name: 130_F_6
-------------- Summary Statistics
    Sequencing vector: M13; M77815; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 183012 bases at least Q40
    Consensus quality: 197888 bases at least Q30
    Consensus quality: 203091 bases at least Q20
    Insert size: 206812; sum-of-contigs
    Quality coverage: 3.6 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1012: contig of 1012 bp in length
*     1013     1112: gap of 100 bp
*     1113     2361: contig of 1249 bp in length
*     2362     2461: gap of 100 bp
*     2462     3665: contig of 1204 bp in length
*     3666     3765: gap of 100 bp
*     3766     5013: contig of 1248 bp in length
*     5014     5113: gap of 100 bp
*     5114     6685: contig of 1572 bp in length
*     6686     6785: gap of 100 bp
*     6786     7988: contig of 1203 bp in length
*     7989     8088: gap of 100 bp
*     8089     9511: contig of 1423 bp in length
*     9512     9611: gap of 100 bp
*     9612    10884: contig of 1273 bp in length
*    10885    10984: gap of 100 bp
*    10985    12222: contig of 1238 bp in length
*    12223    12322: gap of 100 bp
*    12323    13435: contig of 1113 bp in length
*    13436    13535: gap of 100 bp
*    13536    14793: contig of 1258 bp in length
*    14794    14893: gap of 100 bp
*    14894    16074: contig of 1181 bp in length
*    16075    16174: gap of 100 bp
*    16175    17342: contig of 1168 bp in length
*    17343    17442: gap of 100 bp
*    17443    18511: contig of 1069 bp in length
*    18512    18611: gap of 100 bp
*    18612    20151: contig of 1540 bp in length
*    20152    20251: gap of 100 bp
*    20252    21601: contig of 1350 bp in length
*    21602    21701: gap of 100 bp
*    21702    23636: contig of 1935 bp in length
*    23637    23736: gap of 100 bp
*    23737    25405: contig of 1669 bp in length
*    25406    25505: gap of 100 bp
*    25506    26653: contig of 1148 bp in length
*    26654    26753: gap of 100 bp
*    26754    28002: contig of 1249 bp in length
*    28003    28102: gap of 100 bp
*    28103    29661: contig of 1559 bp in length
*    29662    29761: gap of 100 bp
*    29762    31020: contig of 1259 bp in length
*    31021    31120: gap of 100 bp
*    31121    32758: contig of 1638 bp in length
*    32759    32858: gap of 100 bp
*    32859    34433: contig of 1575 bp in length
*    34434    34533: gap of 100 bp
*    34534    36799: contig of 2266 bp in length
*    36800    36899: gap of 100 bp
*    36900    38788: contig of 1889 bp in length
*    38789    38888: gap of 100 bp
*    38889    41490: contig of 2602 bp in length
*    41491    41590: gap of 100 bp
*    41591    43768: contig of 2178 bp in length
*    43769    43868: gap of 100 bp
*    43869    45634: contig of 1766 bp in length
*    45635    45734: gap of 100 bp
*    45735    46330: contig of 596 bp in length
*    46331    46430: gap of 100 bp
*    46431    48111: contig of 1681 bp in length
*    48112    48211: gap of 100 bp
*    48212    50798: contig of 2587 bp in length
*    50799    50898: gap of 100 bp
*    50899    51541: contig of 643 bp in length
*    51542    51641: gap of 100 bp
*    51642    54229: contig of 2588 bp in length
*    54230    54329: gap of 100 bp
*    54330    57592: contig of 3263 bp in length
*    57593    57692: gap of 100 bp
*    57693    60380: contig of 2688 bp in length
*    60381    60480: gap of 100 bp
*    60481    64861: contig of 4381 bp in length
*    64862    64961: gap of 100 bp
*    64962    67169: contig of 2208 bp in length
*    67170    67269: gap of 100 bp
*    67270    70318: contig of 3049 bp in length
*    70319    70418: gap of 100 bp
*    70419    74740: contig of 4322 bp in length
*    74741    74840: gap of 100 bp
*    74841    79050: contig of 4210 bp in length
*    79051    79150: gap of 100 bp
*    79151    84593: contig of 5443 bp in length
*    84594    84693: gap of 100 bp
*    84694    94483: contig of 9790 bp in length
*    94484    94583: gap of 100 bp
*    94584   105653: contig of 11070 bp in length
*   105654   105753: gap of 100 bp
*   105754   116723: contig of 10970 bp in length
*   116724   116823: gap of 100 bp
*   116824   129473: contig of 12650 bp in length
*   129474   129573: gap of 100 bp
*   129574   146634: contig of 17061 bp in length
*   146635   146734: gap of 100 bp
*   146735   162693: contig of 15959 bp in length
*   162694   162793: gap of 100 bp
*   162794   178165: contig of 15372 bp in length
*   178166   178265: gap of 100 bp
*   178266   196145: contig of 17880 bp in length
*   196146   196245: gap of 100 bp
*   196246   211812: contig of 15567 bp in length.
FEATURES             Location/Qualifiers
     source          1. .211812
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"

  Query Match             1.7%;  Score 52;  DB 14;  Length 211812;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211005 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 210954


RESULT 194
AC135988/c
LOCUS       AC135988              213947 bp    DNA     linear   HTG 21-NOV-2002
DEFINITION  Homo sapiens chromosome 15 clone RP11-701O21 map 15, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC135988
VERSION     AC135988.2  GI:25141060
```

```
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 213947)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-701O21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 213947)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 213947)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Nov 21, 2002 this sequence version replaced gi:24414513.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L28438
                Center clone name: 701_O_21
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 213175 bases at least Q40
                Consensus quality: 213476 bases at least Q30
                Consensus quality: 213580 bases at least Q20
                Insert size: 211000; agarose-fp
                Insert size: 213647; sum-of-contigs
                Quality coverage: 13.0 in Q20 bases; agarose-fp
                Quality coverage: 12.8 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11215: contig of 11215 bp in length
            *    11216    11315: gap of 100 bp
            *    11316    44385: contig of 33070 bp in length
            *    44386    44485: gap of 100 bp
            *    44486   112030: contig of 67545 bp in length
            *   112031   112130: gap of 100 bp
            *   112131   213947: contig of 101817 bp in length.
FEATURES             Location/Qualifiers
     source          1. .213947
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-701O21"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .11215
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             11216. .11315
                     /estimated_length=100
     misc_feature    11316. .44385
                     /note="assembly_fragment"
     gap             44386. .44485
                     /estimated_length=100
     misc_feature    44486. .112030
                     /note="assembly_fragment"
     gap             112031. .112130
                     /estimated_length=100
     misc_feature    112131. .213947
                     /note="assembly_fragment"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 213947;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97995 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 97944


RESULT 195
AC135989
LOCUS       AC135989              215780 bp    DNA     linear   HTG 21-NOV-2002
DEFINITION  Homo sapiens chromosome 15 clone RP11-800O12 map 15, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC135989
VERSION     AC135989.2  GI:25141063
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 215780)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-800O12
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 215780)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 215780)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Nov 21, 2002 this sequence version replaced gi:24414514.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L28440
                Center clone name: 800_O_12
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 215088 bases at least Q40
                Consensus quality: 215264 bases at least Q30
                Consensus quality: 215336 bases at least Q20
                Insert size: 215000; agarose-fp
                Insert size: 215480; sum-of-contigs
                Quality coverage: 15.5 in Q20 bases; agarose-fp
                Quality coverage: 15.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    60439: contig of 60439 bp in length
            *    60440    60539: gap of 100 bp
            *    60540    61230: contig of 691 bp in length
            *    61231    61330: gap of 100 bp
            *    61331    61917: contig of 587 bp in length
            *    61918    62017: gap of 100 bp
            *    62018   215780: contig of 153763 bp in length.
FEATURES             Location/Qualifiers
     source          1. .215780
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-800O12"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .60439
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             60440. .60539
                     /estimated_length=100
     misc_feature    60540. .61230
                     /note="assembly_fragment"
     gap             61231. .61330
                     /estimated_length=100
     misc_feature    61331. .61917
                     /note="assembly_fragment"
     gap             61918. .62017
                     /estimated_length=100
     misc_feature    62018. .215780
                     /note="assembly_fragment"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 215780;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6909 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 6960


RESULT 196
AC002429/c
LOCUS       AC002429              234053 bp    DNA     linear   PRI 03-OCT-2003
DEFINITION  Homo sapiens BAC clone GS1-200K5 from 7, complete sequence.
ACCESSION   AC002429
VERSION     AC002429.1  GI:2335067
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 234053)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 234053)
  AUTHORS   Duckels,G., Graves,T. and Hawkins,M.
  TITLE     The sequence of Homo sapiens BAC clone GS1-200K5
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 234053)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1997) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 234053)
```

```
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 234053)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 234053)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 234053)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_GS200K05
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            This clone is from the first BAC library from Genome Systems, Inc.
            (http://www.genomesystems.com).
            Cell line:  lymphoblastoid
            Haplotypes:  two
            VECTOR:  pBeloBAC
            Selection:  chloramphenicol.
FEATURES             Location/Qualifiers
     source          1. .234053
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="GS1-200K5"
                     /clone_lib="GSBAC1"
     repeat_region   31. .81
                     /rpt_family="CT-rich"
     repeat_region   477. .628
                     /rpt_family="MIR"
     repeat_region   2145. .2168
                     /rpt_family="AT_rich"
     repeat_region   3013. .3036
                     /rpt_family="AT_rich"
     repeat_region   3230. .3538
                     /rpt_family="L1"
     repeat_region   3539. .3625
                     /rpt_family="MER2_type"
     repeat_region   3636. .4018
                     /rpt_family="MER2_type"
     repeat_region   4040. .4191
                     /rpt_family="L1"
     repeat_region   7520. .7597
                     /rpt_family="Alu"
     repeat_region   7598. .8047
                     /rpt_family="L1"
     repeat_region   8048. .8335
                     /rpt_family="Alu"
     misc_feature    8073. .8336
                     /note="CpG_island (%GC=60.6, o/e=0.88, #CpGs=21)"
     repeat_region   8336. .8580
                     /rpt_family="L1"
     repeat_region   10511. .10654
                     /rpt_family="Alu"
     repeat_region   10869. .10970
                     /rpt_family="L2"
     repeat_region   13585. .13894
                     /rpt_family="Alu"
     repeat_region   14043. .14102
                     /rpt_family="MER1_type"
     repeat_region   15225. .15261
                     /rpt_family="(T)n"
     repeat_region   16102. .16139
                     /rpt_family="AT_rich"
     repeat_region   16215. .16474
                     /rpt_family="Alu"
     repeat_region   16669. .16826
                     /rpt_family="L2"
     repeat_region   16838. .17493
                     /rpt_family="L2"
     repeat_region   17533. .17630
                     /rpt_family="L2"
     repeat_region   17631. .17938
                     /rpt_family="Alu"
     repeat_region   17939. .18058
                     /rpt_family="L2"
     repeat_region   18676. .18974
                     /rpt_family="Alu"
     repeat_region   19969. .20188
                     /rpt_family="MER1_type"
     repeat_region   20255. .20328
                     /rpt_family="MER1_type"
     repeat_region   20349. .20503
                     /rpt_family="Alu"
     repeat_region   20506. .20572
                     /rpt_family="MER1_type"
     repeat_region   20764. .20907
                     /rpt_family="L2"
     repeat_region   21114. .21414
                     /rpt_family="Alu"
     repeat_region   25771. .25861
                     /rpt_family="CR1"
     repeat_region   26146. .26170
                     /rpt_family="AT_rich"
     repeat_region   26220. .26642
                     /rpt_family="L1"
     repeat_region   27343. .27555
                     /rpt_family="L1"
     repeat_region   28645. .28672
                     /rpt_family="(CATATA)n"
     repeat_region   28916. .29059
                     /rpt_family="MER1_type"
     repeat_region   29080. .29472
                     /rpt_family="ERV1"
     repeat_region   29524. .29577
```

```
                     /rpt_family="MER1_type"
     repeat_region   30140. .30236
                     /rpt_family="MIR"
     repeat_region   30312. .31450
                     /rpt_family="L1"
     repeat_region   31525. .31829
                     /rpt_family="Alu"
     repeat_region   31873. .32741
                     /rpt_family="L1"
     repeat_region   32761. .33823
                     /rpt_family="L1"
     repeat_region   33825. .34218
                     /rpt_family="ERV1"
     repeat_region   34230. .34299
                     /rpt_family="ERV1"
     repeat_region   34304. .34494
                     /rpt_family="L1"
     repeat_region   34495. .34789
                     /rpt_family="Alu"
     repeat_region   34790. .35979
                     /rpt_family="L1"
     repeat_region   35977. .36335
                     /rpt_family="L1"
     repeat_region   36342. .36677
                     /rpt_family="L1"
     repeat_region   36678. .36981
                     /rpt_family="Alu"
     repeat_region   36982. .37759
                     /rpt_family="L1"

  Query Match             1.7%;  Score 52;  DB 8;  Length 234053;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     223398 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 223347


RESULT 197
AC145896/c
LOCUS       AC145896               256000 bp    DNA     linear   HTG 13-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-8H23, *** SEQUENCING IN
            PROGRESS ***, 48 unordered pieces.
ACCESSION   AC145896
VERSION     AC145896.1  GI:33386828
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 256000)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 256000)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 256000)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: C_PT008H23
            ------------------------------------------------
            ------------- Summary Statistics --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 237274 bases at least Q40
            Consensus quality: 240505 bases at least Q30
            Consensus quality: 242221 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 48 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1185: contig of 1185 bp in length
            *     1186     1285: gap of unknown length
            *     1286     2469: contig of 1184 bp in length
            *     2470     2569: gap of unknown length
            *     2570     3979: contig of 1410 bp in length
            *     3980     4079: gap of unknown length
            *     4080     5295: contig of 1216 bp in length
            *     5296     5395: gap of unknown length
            *     5396     6630: contig of 1235 bp in length
            *     6631     6730: gap of unknown length
            *     6731     7920: contig of 1190 bp in length
            *     7921     8020: gap of unknown length
            *     8021     9452: contig of 1432 bp in length
            *     9453     9552: gap of unknown length
            *     9553    10805: contig of 1253 bp in length
            *    10806    10905: gap of unknown length
            *    10906    12145: contig of 1240 bp in length
            *    12146    12245: gap of unknown length
            *    12246    13423: contig of 1178 bp in length
            *    13424    13523: gap of unknown length
            *    13524    15026: contig of 1503 bp in length
            *    15027    15126: gap of unknown length
            *    15127    16408: contig of 1282 bp in length
            *    16409    16508: gap of unknown length
            *    16509    17900: contig of 1392 bp in length
            *    17901    18000: gap of unknown length
            *    18001    19931: contig of 1931 bp in length
            *    19932    20031: gap of unknown length
            *    20032    21219: contig of 1188 bp in length
            *    21220    21319: gap of unknown length
            *    21320    23430: contig of 2111 bp in length
            *    23431    23530: gap of unknown length
            *    23531    25074: contig of 1544 bp in length
            *    25075    25174: gap of unknown length
            *    25175    26544: contig of 1370 bp in length
            *    26545    26644: gap of unknown length
            *    26645    28429: contig of 1785 bp in length
            *    28430    28529: gap of unknown length
            *    28530    29909: contig of 1380 bp in length
            *    29910    30009: gap of unknown length
            *    30010    31609: contig of 1600 bp in length
            *    31610    31709: gap of unknown length
            *    31710    34226: contig of 2517 bp in length
            *    34227    34326: gap of unknown length
            *    34327    35499: contig of 1173 bp in length
            *    35500    35599: gap of unknown length
            *    35600    36828: contig of 1229 bp in length
            *    36829    36928: gap of unknown length
            *    36929    39530: contig of 2602 bp in length
            *    39531    39630: gap of unknown length
            *    39631    40715: contig of 1085 bp in length
            *    40716    40815: gap of unknown length
            *    40816    42410: contig of 1595 bp in length
            *    42411    42510: gap of unknown length
```

```
            *    42511    43677: contig of 1167 bp in length
            *    43678    43777: gap of unknown length
            *    43778    44933: contig of 1156 bp in length
            *    44934    45033: gap of unknown length
            *    45034    46566: contig of 1533 bp in length
            *    46567    46666: gap of unknown length
            *    46667    48332: contig of 1666 bp in length
            *    48333    48432: gap of unknown length
            *    48433    50446: contig of 2014 bp in length
            *    50447    50546: gap of unknown length
            *    50547    51956: contig of 1410 bp in length
            *    51957    52056: gap of unknown length
            *    52057    54338: contig of 2282 bp in length
            *    54339    54438: gap of unknown length
            *    54439    56408: contig of 1970 bp in length
            *    56409    56508: gap of unknown length
            *    56509    60179: contig of 3671 bp in length
            *    60180    60279: gap of unknown length
            *    60280    64269: contig of 3990 bp in length
            *    64270    64369: gap of unknown length
            *    64370    65819: contig of 1450 bp in length
            *    65820    65919: gap of unknown length
            *    65920    70603: contig of 4684 bp in length
            *    70604    70703: gap of unknown length
            *    70704    74447: contig of 3744 bp in length
            *    74448    74547: gap of unknown length
            *    74548    80301: contig of 5754 bp in length
            *    80302    80401: gap of unknown length
            *    80402    93561: contig of 13160 bp in length
            *    93562    93661: gap of unknown length
            *    93662   107828: contig of 14167 bp in length
            *   107829   107928: gap of unknown length
            *   107929   129690: contig of 21762 bp in length
            *   129691   129790: gap of unknown length
            *   129791   150182: contig of 20392 bp in length
            *   150183   150282: gap of unknown length
            *   150283   181315: contig of 31033 bp in length
            *   181316   181415: gap of unknown length
            *   181416   207034: contig of 25619 bp in length
            *   207035   207134: gap of unknown length
            *   207135   256000: contig of 48866 bp in length.
FEATURES             Location/Qualifiers
     source          1. .256000
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-8H23"
     misc_feature    1. .1185
                     /note="assembly_name:Contig23"
     gap             1186. .1285
                     /estimated_length=unknown
     misc_feature    1286. .2469
                     /note="assembly_name:Contig25"
     gap             2470. .2569
                     /estimated_length=unknown
     misc_feature    2570. .3979
                     /note="assembly_name:Contig31"
     gap             3980. .4079
                     /estimated_length=unknown
     misc_feature    4080. .5295
                     /note="assembly_name:Contig33"
     gap             5296. .5395
                     /estimated_length=unknown
     misc_feature    5396. .6630
                     /note="assembly_name:Contig37"
     gap             6631. .6730
                     /estimated_length=unknown
     misc_feature    6731. .7920
                     /note="assembly_name:Contig38"
     gap             7921. .8020
                     /estimated_length=unknown
     misc_feature    8021. .9452
                     /note="assembly_name:Contig39"
     gap             9453. .9552
                     /estimated_length=unknown
     misc_feature    9553. .10805
                     /note="assembly_name:Contig41"
     gap             10806. .10905
                     /estimated_length=unknown
     misc_feature    10906. .12145
                     /note="assembly_name:Contig43"
     gap             12146. .12245
                     /estimated_length=unknown
     misc_feature    12246. .13423
                     /note="assembly_name:Contig44"
     gap             13424. .13523
                     /estimated_length=unknown
     misc_feature    13524. .15026
                     /note="assembly_name:Contig51"
     gap             15027. .15126
                     /estimated_length=unknown
     misc_feature    15127. .16408
                     /note="assembly_name:Contig52"
     gap             16409. .16508

  Query Match             1.7%;  Score 52;  DB 14;  Length 256000;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175847 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 175796


RESULT 198
AC145065
LOCUS       AC145065               279430 bp    DNA     linear   HTG 06-JUL-2005
DEFINITION  Pan troglodytes clone rp43-47k21, WORKING DRAFT SEQUENCE, 19
            unordered pieces.
ACCESSION   AC145065
VERSION     AC145065.23  GI:68989303
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 279430)
  AUTHORS   Yu,J., Do,T. and Roe,B.A.
  TITLE     Pan troglodytes BAC Clone rp43-47k21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 279430)
  AUTHORS   Yu,J., Do,T. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 279430)
  AUTHORS   Yu,J., Do,T. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2005) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Jul 6, 2005 this sequence version replaced gi:68342157.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
```

```
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     2054: contig of 2054 bp in length
          *     2055     2154: gap of unknown length
          *     2155     5485: contig of 3331 bp in length
          *     5486     5585: gap of unknown length
          *     5586     8523: contig of 2938 bp in length
          *     8524     8623: gap of unknown length
          *     8624    11066: contig of 2443 bp in length
          *    11067    11166: gap of unknown length
          *    11167    14580: contig of 3414 bp in length
          *    14581    14680: gap of unknown length
          *    14681    17963: contig of 3283 bp in length
          *    17964    18063: gap of unknown length
          *    18064    21381: contig of 3318 bp in length
          *    21382    21481: gap of unknown length
          *    21482    25967: contig of 4486 bp in length
          *    25968    26067: gap of unknown length
          *    26068    30350: contig of 4283 bp in length
          *    30351    30450: gap of unknown length
          *    30451    37058: contig of 6608 bp in length
          *    37059    37158: gap of unknown length
          *    37159    44120: contig of 6962 bp in length
          *    44121    44220: gap of unknown length
          *    44221    56124: contig of 11904 bp in length
          *    56125    56224: gap of unknown length
          *    56225    70495: contig of 14271 bp in length
          *    70496    70595: gap of unknown length
          *    70596    88931: contig of 18336 bp in length
          *    88932    89031: gap of unknown length
          *    89032   106819: contig of 17788 bp in length
          *   106820   106919: gap of unknown length
          *   106920   140897: contig of 33978 bp in length
          *   140898   140997: gap of unknown length
          *   140998   178266: contig of 37269 bp in length
          *   178267   178366: gap of unknown length
          *   178367   215146: contig of 36780 bp in length
          *   215147   215246: gap of unknown length
          *   215247   279430: contig of 64184 bp in length.
FEATURES             Location/Qualifiers
     source          1. .279430
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="rp43-47k21"
                     /clone_lib="RPCI - 43 Male Chimpanzee BAC Library"
     gap             2055. .2154
                     /estimated_length=unknown
     gap             5486. .5585
                     /estimated_length=unknown
     gap             8524. .8623
                     /estimated_length=unknown
     gap             11067. .11166
                     /estimated_length=unknown
     gap             14581. .14680
                     /estimated_length=unknown
     gap             17964. .18063
                     /estimated_length=unknown
     gap             21382. .21481
                     /estimated_length=unknown
     gap             25968. .26067
                     /estimated_length=unknown
     gap             30351. .30450
                     /estimated_length=unknown
     gap             37059. .37158
                     /estimated_length=unknown
     gap             44121. .44220
                     /estimated_length=unknown
     gap             56125. .56224
                     /estimated_length=unknown
     gap             70496. .70595
                     /estimated_length=unknown
     gap             88932. .89031
                     /estimated_length=unknown
     gap             106820. .106919
                     /estimated_length=unknown
     gap             140898. .140997
                     /estimated_length=unknown
     gap             178267. .178366
                     /estimated_length=unknown
     gap             215147. .215246
                     /estimated_length=unknown
ORIGIN

  Query Match             1.7%;  Score 52;  DB 14;  Length 279430;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204125 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 204176


RESULT 199
CS039416
LOCUS       CS039416              349980 bp    DNA     linear   PAT 22-MAR-2005
DEFINITION  Sequence 226 from Patent WO2005019477.
ACCESSION   CS039416
VERSION     CS039416.1  GI:61847423
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lewin,J., Berlin,K., Hildmann,T., Olek,A., Beck,S. and Novik,K.
  TITLE     Methods and compositions for differentiating tissues or cell types
            using epigenetic markers
  JOURNAL   Patent: WO 2005019477-A 226 03-MAR-2005;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="unknown base
                     seq 205 to long, 4.267.840
                     replaced by
                     new seq 205, from 0.000.001 to 0.349.980
                     new seq 222, from 0.300.001 to 0.649.980
                     new seq 223, from 0.600.001 to 0.949.980
                     new seq 224, from 0.900.001 to 1.249.980
                     new seq 225, from 1.200.001 to 1.549.980
                     new seq 226, from 1.500.001 to 1.849.980
                     new seq 227, from 1.800.001 to 2.149.980
                     new seq 228, from 2.100.001 to 2.449.980
                     new seq 229, from 2.400.001 to 2.749.980
                     new seq 230, from 2.700.001 to 3.049.980
                     new seq 231, from 3.000.001 to 3.349.980
                     new seq 232, from 3.300.001 to 3.649.980
                     new seq 233, from 3.600.001 to 3.949.980
                     new seq 234, from 3.900.001 to 4.249.980
                     new seq 235, from 4.200.001 to 4.549.980
                     new seq 236, from 4.500.001 to 4.647.455"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 6;  Length 349980;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273885 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 273936
```

```
RESULT 200
BV634964/c
LOCUS       BV634964                715 bp    DNA     linear   STS 15-APR-2005
DEFINITION  S217P6621FB7.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV634964
VERSION     BV634964.1  GI:62625142
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 715
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>715
ORIGIN
```

```
  Query Match             1.6%;  Score 51;  DB 10;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3072 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 430


Search completed: May 11, 2006, 08:43:47
Job time : 15448 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 04:27:18 ; Search time 11817 Seconds
(without alignments)
12360.944 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 1

Total number of hits satisfying chosen parameters: 82156240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	646	20.7	660	5	BX116028	BX116028 BX116028
	2	613	19.6	670	6	CF145408	CF145408 UI-HF-CB0
	3	598	19.2	677	6	CF145448	CF145448 UI-HF-CB0
c	4	531	17.0	685	3	BM671616	BM671616 UI-E-CQ1-
	5	519	16.6	578	3	BM707056	BM707056 UI-E-CR1-
	6	500	16.0	500	3	BM710194	BM710194 UI-E-CQ1-
	7	450	14.4	655	3	BQ187216	BQ187216 UI-E-EJ1-
c	8	427	13.7	476	1	AW302149	AW302149 xs01f06.x
c	9	413	13.2	455	1	AI375213	AI375213 tc10f06.x
c	10	412	13.2	480	3	BM673230	BM673230 UI-E-CR1-
	11	395	12.7	972	5	BQ682843	BQ682843 AGENCOURT
	12	348	11.1	1376	3	BM562937	BM562937 AGENCOURT
	13	333	10.7	595	8	W81663	W81663 zd85e06.r1
c	14	290	9.3	605	3	BI458731	BI458731 603199879
	15	180	5.8	728	5	BX645932	BX645932 DKFZp781B
c	16	144	4.6	449	8	W81222	W81222 zd85e06.s1
	17	53	1.7	709	10	AG013685	AG013685 Homo sapi
	18	53	1.7	709	10	AG013687	AG013687 Homo sapi
	19	53	1.7	718	10	AG013686	AG013686 Homo sapi
c	20	52	1.7	354	1	AA679936	AA679936 ac85g01.s
c	21	52	1.7	441	1	AI732120	AI732120 ac86a01.x
c	22	52	1.7	441	1	AI732180	AI732180 ac85g01.x
	23	52	1.7	520	7	CN277748	CN277748 170006001
c	24	52	1.7	529	1	AA680243	AA680243 ac86a01.s
c	25	52	1.7	553	7	CK825142	CK825142 ig44h08.y
	26	52	1.7	555	7	CK825141	CK825141 ig44h08.x
	27	52	1.7	568	9	AQ526715	AQ526715 HS_5216_A
c	28	51	1.6	373	9	BZ606000	BZ606000 WHAAS80TR
	29	50	1.6	350	7	CR545224	CR545224 DKFZp459E
	30	50	1.6	471	6	CB140816	CB140816 K-EST0194
c	31	50	1.6	572	9	AQ487295	AQ487295 RPCI-11-2
c	32	50	1.6	634	9	B47416	B47416 CIT978SK-16
	33	50	1.6	656	10	AG145253	AG145253 Pan trogl
	34	50	1.6	694	1	AV734315	AV734315 AV734315
	35	50	1.6	706	7	CN413517	CN413517 170005326
	36	50	1.6	770	6	CD357076	CD357076 AGENCOURT
	37	49	1.6	797	9	AY758761	AY758761 RP43-99b9
c	38	48	1.5	452	1	AW006453	AW006453 wr16h01.x
c	39	48	1.5	548	9	AQ757333	AQ757333 HS_5484_A
c	40	47	1.5	457	9	AQ121008	AQ121008 HS_3073_A
c	41	47	1.5	470	9	AQ029209	AQ029209 RPCI11-41
c	42	47	1.5	566	9	AQ508783	AQ508783 RPCI-11-2
c	43	47	1.5	847	2	BF526343	BF526343 602070608
c	44	46	1.5	216	1	AA078268	AA078268 7H48E05 C
	45	46	1.5	255	1	AI309360	AI309360 tb51c03.x
	46	46	1.5	313	5	BX475355	BX475355 DKFZp686M
	47	46	1.5	314	1	AW841960	AW841960 IL5-CN002
c	48	46	1.5	319	6	CD514740	CD514740 AGENCOURT
c	49	46	1.5	329	1	AA661680	AA661680 nu87b06.s
c	50	46	1.5	341	2	BE061760	BE061760 RC1-BT025
	51	46	1.5	372	2	BF767274	BF767274 RC2-CN008
c	52	46	1.5	387	9	AQ201612	AQ201612 RPCI11-59
c	53	46	1.5	389	9	AQ134220	AQ134220 HS_3047_B
c	54	46	1.5	405	1	AA503019	AA503019 ng20d04.s
	55	46	1.5	406	9	AQ284177	AQ284177 RPCI11-80
c	56	46	1.5	445	1	AW815323	AW815323 QV0-ST021
	57	46	1.5	462	1	AA504863	AA504863 ab03b02.s
c	58	46	1.5	464	1	AA128941	AA128941 zo11c09.s
	59	46	1.5	482	9	AQ383997	AQ383997 RPCI11-15
	60	46	1.5	507	9	AQ132449	AQ132449 HS_3003_A
c	61	46	1.5	527	5	BX641230	BX641230 DKFZp686D
	62	46	1.5	550	5	BU960416	BU960416 AGENCOURT
c	63	46	1.5	564	11	CR960770	CR960770 Homo sapi
	64	46	1.5	574	7	CN265832	CN265832 170004248
c	65	46	1.5	587	9	B95171	B95171 CIT-HSP-217
c	66	46	1.5	638	9	AQ017813	AQ017813 CIT-HSP-2
	67	46	1.5	648	5	BX956845	BX956845 DKFZp781H
c	68	46	1.5	657	5	BX641474	BX641474 DKFZp686I
	69	46	1.5	687	6	CA430010	CA430010 UI-H-FH1-
	70	45	1.4	258	5	BU564032	BU564032 AGENCOURT
	71	45	1.4	287	5	BX953709	BX953709 DKFZp781C
c	72	45	1.4	309	9	AQ485964	AQ485964 RPCI-11-2
	73	45	1.4	352	9	AQ381523	AQ381523 RPCI11-16
c	74	45	1.4	408	9	AQ485959	AQ485959 RPCI-11-2
c	75	45	1.4	415	5	BU951869	BU951869 in74a04.x
c	76	45	1.4	417	9	AQ065412	AQ065412 HS_2224_B
c	77	45	1.4	424	1	AA085683	AA085683 zn53d01.s
c	78	45	1.4	436	6	CB132267	CB132267 K-EST0182
	79	45	1.4	439	9	AQ009550	AQ009550 CIT-HSP-2
c	80	45	1.4	447	9	AQ554616	AQ554616 RPCI-11-4
	81	45	1.4	452	9	AQ170811	AQ170811 HS_3071_A
	82	45	1.4	469	2	BG944968	BG944968 ax57a12.x
c	83	45	1.4	472	8	H63465	H63465 yr53e06.r1
	84	45	1.4	494	5	BX954324	BX954324 DKFZp781B
c	85	45	1.4	503	9	AQ038737	AQ038737 CIT-HSP-2
c	86	45	1.4	529	6	CB142858	CB142858 K-EST0196
c	87	45	1.4	529	9	AQ333702	AQ333702 HS_5013_A
c	88	45	1.4	568	1	AU158612	AU158612 AU158612
c	89	45	1.4	599	3	BP872324	BP872324 BP872324
	90	45	1.4	627	10	AG071092	AG071092 Pan trogl
c	91	45	1.4	640	10	AG116078	AG116078 Pan trogl
c	92	45	1.4	641	9	AQ540193	AQ540193 RPCI-11-3
c	93	45	1.4	680	10	AG174909	AG174909 Pan trogl
	94	45	1.4	712	7	CN274316	CN274316 170006000
	95	45	1.4	722	9	AQ349458	AQ349458 RPCI11-11

	96	45	1.4	726	10	AG009127 AG009127 Homo sapi
c	97	45	1.4	731	10	AG009138 AG009138 Homo sapi
	98	45	1.4	741	10	AG009129 AG009129 Homo sapi
	99	45	1.4	754	5	BX457023 BX457023 BX457023
c	100	45	1.4	768	2	BG260049 BG260049 602371633
c	101	45	1.4	781	5	BX411269 BX411269 BX411269
	102	45	1.4	830	6	CD242479 CD242479 AGENCOURT
c	103	45	1.4	856	5	BU959380 BU959380 AGENCOURT
c	104	45	1.4	874	9	BZ771376 BZ771376 mcs76c11.
c	105	45	1.4	948	6	CF596843 CF596843 AGENCOURT
c	106	45	1.4	988	5	BQ896885 BQ896885 AGENCOURT
	107	45	1.4	1123	5	BX404721 BX404721 BX404721
	108	44	1.4	203	2	BG959135 BG959135 PM4-CT080
c	109	44	1.4	219	2	BG010132 BG010132 MR3-GN018
	110	44	1.4	281	2	BF893386 BF893386 QV3-MT012
	111	44	1.4	281	9	BH609712 BH609712 HIV18E03
c	112	44	1.4	294	1	AA493894 AA493894 nh07b12.s
c	113	44	1.4	301	1	AI382825 AI382825 ta72h10.x
	114	44	1.4	301	2	BF879334 BF879334 IL3-ET011
c	115	44	1.4	325	2	BF914587 BF914587 IL3-UT011
c	116	44	1.4	337	9	AQ059204 AQ059204 CIT-HSP-2
c	117	44	1.4	338	8	N63149 N63149 yz37e10.s1
c	118	44	1.4	352	9	AQ092587 AQ092587 HS_3003_B
c	119	44	1.4	353	1	AI367551 AI367551 qv99e05.x
c	120	44	1.4	365	1	AA654781 AA654781 nt73g04.s
	121	44	1.4	377	9	AQ061311 AQ061311 CIT-HSP-2
c	122	44	1.4	378	1	AA653916 AA653916 nt80a05.s
c	123	44	1.4	381	9	AQ239365 AQ239365 CIT-HSP-2
	124	44	1.4	382	5	BU561388 BU561388 AGENCOURT
c	125	44	1.4	389	1	AA728990 AA728990 nw22g06.s
	126	44	1.4	409	8	H44630 H44630 yp19g10.s1
c	127	44	1.4	412	9	AQ605062 AQ605062 HS_2119_B
c	128	44	1.4	421	7	CR769029 CR769029 DKFZp468I
	129	44	1.4	426	8	N20066 N20066 yx28g04.s1
c	130	44	1.4	435	7	CR546695 CR546695 DKFZp470L
c	131	44	1.4	438	9	B60059 B60059 CIT-HSP-385
	132	44	1.4	447	1	AL713064 AL713064 DKFZp686C
	133	44	1.4	448	1	AI469968 AI469968 tj89c03.x
c	134	44	1.4	448	9	AQ705037 AQ705037 HS_5521_B
c	135	44	1.4	466	1	AU157238 AU157238 AU157238
	136	44	1.4	466	5	BX111733 BX111733 BX111733
c	137	44	1.4	468	2	BE674703 BE674703 7e94e04.x
c	138	44	1.4	472	8	T84449 T84449 yd32d04.r1
c	139	44	1.4	473	3	BM983814 BM983814 UI-CF-DU1
	140	44	1.4	483	3	BM314257 BM314257 ig52c03.x
c	141	44	1.4	484	9	AQ276534 AQ276534 CITBI-E1-
	142	44	1.4	485	5	BQ432539 BQ432539 AGENCOURT
	143	44	1.4	486	8	N35896 N35896 yy28c08.s1
	144	44	1.4	494	1	AI908381 AI908381 RC-BT170-
c	145	44	1.4	495	8	R86218 R86218 yp88a10.r1
	146	44	1.4	512	2	BE146359 BE146359 MR0-HT020
	147	44	1.4	516	1	AI888752 AI888752 wn36a07.x
c	148	44	1.4	517	2	BE940436 BE940436 RC3-UT006
	149	44	1.4	527	9	AQ026637 AQ026637 CIT-HSP-2
c	150	44	1.4	530	7	CN389220 CN389220 170006003
	151	44	1.4	538	6	CD691041 CD691041 EST7564 h
c	152	44	1.4	540	7	CV418654 CV418654 RC3-UT006
c	153	44	1.4	543	9	B55060 B55060 CIT-HSP-385
c	154	44	1.4	543	9	AQ713671 AQ713671 HS_5399_B
c	155	44	1.4	551	1	AA715817 AA715817 nw25c05.s
c	156	44	1.4	558	6	CB129216 CB129216 K-EST0178
	157	44	1.4	558	7	CV703188 CV703188 TgESTzys3
	158	44	1.4	563	1	AW877914 AW877914 MR3-OT000
	159	44	1.4	569	9	AQ591303 AQ591303 HS_2122_B
c	160	44	1.4	581	6	CD723228 CD723228 oj19e04.y
c	161	44	1.4	582	6	CB218356 CB218356 NISC_nb08
c	162	44	1.4	590	1	AA206418 AA206418 zq51a05.s
	163	44	1.4	602	9	B70450 B70450 CIT-HSP-205
	164	44	1.4	614	1	AL691931 AL691931 DKFZp313B
	165	44	1.4	614	10	CL246302 CL246302 HSC 01025
c	166	44	1.4	621	1	AL120523 AL120523 DKFZp761O
	167	44	1.4	627	7	CN267712 CN267712 170005316
c	168	44	1.4	633	9	AQ542374 AQ542374 RPCI-11-3
	169	44	1.4	637	5	BX482192 BX482192 DKFZp686H
c	170	44	1.4	638	9	AQ697816 AQ697816 HS_5536_B
	171	44	1.4	641	1	AV733586 AV733586 AV733586
c	172	44	1.4	649	10	AG025905 AG025905 Homo sapi
	173	44	1.4	650	10	AG080818 AG080818 Pan trogl
	174	44	1.4	652	1	AV722027 AV722027 AV722027
c	175	44	1.4	654	10	AG106557 AG106557 Pan trogl
	176	44	1.4	655	1	AW959999 AW959999 EST372070
	177	44	1.4	694	9	AQ740363 AQ740363 HS_5501_A
c	178	44	1.4	696	5	BU616112 BU616112 UI-H-DF0-
	179	44	1.4	706	9	AQ262202 AQ262202 CITBI-E1-
	180	44	1.4	715	5	BX504396 BX504396 DKFZp686O
c	181	44	1.4	716	5	BX098422 BX098422 BX098422
c	182	44	1.4	745	10	AG175100 AG175100 Pan trogl
	183	44	1.4	781	6	CB308313 CB308313 AGENCOURT
c	184	44	1.4	803	9	BZ610701 BZ610701 WHAAJ70TR
	185	44	1.4	828	6	CA442440 CA442440 UI-H-DP0-
c	186	44	1.4	830	6	CB307968 CB307968 AGENCOURT
c	187	44	1.4	830	9	AQ781442 AQ781442 HS_3117_A
c	188	44	1.4	834	9	BZ612098 BZ612098 WHADP72TR
c	189	44	1.4	847	9	AQ781550 AQ781550 HS_3176_A
	190	44	1.4	867	2	BF573285 BF573285 602079678
c	191	44	1.4	903	9	AQ749032 AQ749032 HS_5575_A
	192	44	1.4	982	1	AL570050 AL570050 AL570050
c	193	44	1.4	1527	4	CR622010 CR622010 full-leng
c	194	43	1.4	108	2	BE141837 BE141837 IL5-HT011
	195	43	1.4	132	1	AA728876 AA728876 nv37c08.r
c	196	43	1.4	132	2	BG012771 BG012771 IL5-GN023
c	197	43	1.4	151	7	CR543084 CR543084 DKFZp470G
c	198	43	1.4	154	8	N73776 N73776 za61h08.s1
	199	43	1.4	157	1	AI611895 AI611895 tt59f11.x
c	200	43	1.4	168	1	AA018823 AA018823 ze57e09.r
c	201	43	1.4	196	1	AA478297 AA478297 zu45d06.s
	202	43	1.4	207	8	F17774 F17774 HSPD02335 H
c	203	43	1.4	213	2	BG616769 BG616769 602614995
	204	43	1.4	224	7	CN278030 CN278030 170006001
c	205	43	1.4	238	1	AW818514 AW818514 RC1-ST027
c	206	43	1.4	246	1	AI886467 AI886467 wn39c10.x
c	207	43	1.4	252	9	AQ286384 AQ286384 RPCI11-79
	208	43	1.4	254	1	AA501777 AA501777 ng05d09.s
c	209	43	1.4	261	2	BG151598 BG151598 nag61h12.
	210	43	1.4	265	1	AA478355 AA478355 zu45d06.r
c	211	43	1.4	277	1	AW848108 AW848108 IL3-CT021
c	212	43	1.4	282	9	AQ378964 AQ378964 RPCI11-16
c	213	43	1.4	294	9	B98981 B98981 CIT-HSP-228
c	214	43	1.4	308	1	AI873852 AI873852 wm42e07.x
c	215	43	1.4	329	7	CN263688 CN263688 170004241
c	216	43	1.4	339	2	BF184537 BF184537 601842843
	217	43	1.4	341	1	AA714110 AA714110 nw05b09.s
	218	43	1.4	341	3	BM709383 BM709383 UI-E-CQ1-
	219	43	1.4	342	7	CK816152 CK816152 AGENCOURT
c	220	43	1.4	345	1	AI355986 AI355986 qy51g05.x
	221	43	1.4	347	8	R76463 R76463 yi63h04.s1
c	222	43	1.4	348	5	BU732762 BU732762 UI-E-CQ1-
	223	43	1.4	350	1	AV734149 AV734149 AV734149
c	224	43	1.4	356	9	AQ263477 AQ263477 CITBI-E1-
c	225	43	1.4	368	1	AA834798 AA834798 od98h01.s
c	226	43	1.4	368	3	BM667130 BM667130 UI-E-DX0-
c	227	43	1.4	370	8	T80208 T80208 yd03a02.r1
c	228	43	1.4	374	7	CV326131 CV326131 CM4-NN000
	229	43	1.4	378	1	AV688904 AV688904 AV688904
c	230	43	1.4	381	1	AA600335 AA600335 ag04e01.s
c	231	43	1.4	383	7	CR737743 CR737743 CR737743
c	232	43	1.4	383	9	AQ062686 AQ062686 CIT-HSP-2
c	233	43	1.4	391	1	AI633479 AI633479 th62e11.x
c	234	43	1.4	397	1	AI336758 AI336758 qw93c01.x
c	235	43	1.4	399	9	AQ312641 AQ312641 RPCI11-95
c	236	43	1.4	400	1	AA633017 AA633017 nq09d08.s
c	237	43	1.4	400	6	CD640343 CD640343 AGENCOURT
c	238	43	1.4	400	8	F01045 F01045 HSB96B042 S
	239	43	1.4	407	9	AQ202847 AQ202847 RPCI11-43
c	240	43	1.4	412	1	AA962276 AA962276 oo53a02.s
c	241	43	1.4	412	1	AI828710 AI828710 tr12d11.x

c 242	43	1.4	420	5	BU074474	BU074474	im76c02.x
243	43	1.4	423	1	AW819125	AW819125	RC3-ST028
c 244	43	1.4	426	8	R20234	R20234	yg18h12.r1
c 245	43	1.4	428	1	AA829110	AA829110	od78a11.s
c 246'	43	1.4	447	1	AA706628	AA706628	ah26d10.s
247	43	1.4	450	2	BE145089	BE145089	CM3-HT019
248	43	1.4	459	1	AL121207	AL121207	DKFZp762C
c 249	43	1.4	459	1	AL121235	AL121235	DKFZp762F
c 250	43	1.4	462	6	CD370252	CD370252	UI-H-FT1-
c 251	43	1.4	465	1	AW675725	AW675725	ba53e09.x
c 252	43	1.4	478	1	AI627168	AI627168	ty60a05.x
c 253	43	1.4	479	8	H14566	H14566	yl25h12.r1
c 254	43	1.4	483	11	CR959520	CR959520	Homo sapi
c 255	43	1.4	489	1	AL119649	AL119649	DKFZp761K
256	43	1.4	493	2	BF854090	BF854090	MR2-EN009
c 257	43	1.4	501	7	CR546841	CR546841	DKFZp470B
c 258	43	1.4	502	9	AQ726636	AQ726636	HS_5411_B
c 259	43	1.4	504	9	AQ135921	AQ135921	HS_3060_A
c 260	43	1.4	505	9	AQ183779	AQ183779	HS_3199_B
c 261	43	1.4	510	2	BG613206	BG613206	602640756
c 262	43	1.4	511	5	BX953870	BX953870	DKFZp781A
c 263	43	1.4	524	9	AQ734876	AQ734876	HS_3051_A
264	43	1.4	527	3	BM770275	BM770275	K-EST0053
c 265	43	1.4	528	9	AQ472287	AQ472287	CITBI-E1-
266	43	1.4	529	7	CN388806	CN388806	170006001
c 267	43	1.4	529	9	AQ460087	AQ460087	HS_5123_B
c 268	43	1.4	532	1	AL698441	AL698441	DKFZp686K
269	43	1.4	543	2	BG254818	BG254818	602369246
270	43	1.4	544	5	BU076635	BU076635	im50d01.y
c 271	43	1.4	547	7	CV571317	CV571317	oe11e07.y
c 272	43	1.4	550	1	AU152518	AU152518	AU152518
c 273	43	1.4	550	1	AW003867	AW003867	ws61d02.x
274	43	1.4	550	5	BX501587	BX501587	DKFZp779N
275	43	1.4	551	5	BU074758	BU074758	im76c02.y
c 276	43	1.4	553	1	AL699840	AL699840	DKFZp686M
c 277	43	1.4	554	1	AW979058	AW979058	EST391168
278	43	1.4	574	2	BE395458	BE395458	601309981
279	43	1.4	576	1	AA601216	AA601216	no14d11.s
c 280	43	1.4	581	1	AW605171	AW605171	QV0-DT002
c 281	43	1.4	583	7	CN263464	CN263464	170004240
c 282	43	1.4	583	9	AZ519883	AZ519883	RPCI-11-7
283	43	1.4	585	6	CD700137	CD700137	EST16661
284	43	1.4	590	3	BM751907	BM751907	K-EST0028
c 285	43	1.4	596	3	BQ010300	BQ010300	UI-H-ED0-
286	43	1.4	606	5	BX486604	BX486604	DKFZp686K
c 287	43	1.4	607	2	BG924293	BG924293	HNC26-1-C
c 288	43	1.4	614	5	BX643423	BX643423	DKFZp781K
c 289	43	1.4	617	5	BX506068	BX506068	DKFZp686D
290	43	1.4	620	2	BE867617	BE867617	601443133
c 291	43	1.4	625	5	BX507175	BX507175	DKFZp779C
c 292	43	1.4	629	6	CD640871	CD640871	AGENCOURT
c 293	43	1.4	630	1	AI064786	AI064786	HA0564 Hu
294	43	1.4	636	10	AG174345	AG174345	Pan trogl
295	43	1.4	638	10	AG102516	AG102516	Pan trogl
c 296	43	1.4	639	9	AQ377619	AQ377619	RPCI11-16
c 297	43	1.4	640	1	AA012964	AA012964	ze27c03.r
c 298	43	1.4	646	10	AG007923	AG007923	Homo sapi
c 299	43	1.4	647	3	BQ019639	BQ019639	UI-H-ED0-
c 300	43	1.4	650	10	AG007924	AG007924	Homo sapi
301	43	1.4	653	10	AG061520	AG061520	Pan trogl
c 302	43	1.4	656	3	BI334807	BI334807	602998923
303	43	1.4	659	3	BI858506	BI858506	603391746
304	43	1.4	660	3	BM742097	BM742097	K-EST0014
305	43	1.4	668	3	BM771614	BM771614	K-EST0055
c 306	43	1.4	669	6	CD368504	CD368504	UI-H-FT1-
c 307	43	1.4	672	9	AQ348932	AQ348932	RPCI11-12
308	43	1.4	677	10	AG158069	AG158069	Pan trogl
309	43	1.4	683	10	AG155728	AG155728	Pan trogl
310	43	1.4	689	3	BM040803	BM040803	603614269
c 311	43	1.4	691	3	BM990097	BM990097	UI-H-DI0-
c 312	43	1.4	692	9	AQ491216	AQ491216	RPCI-11-2
313	43	1.4	692	10	AG141656	AG141656	Pan trogl
c 314	43	1.4	701	3	BI333819	BI333819	602999323
c 315	43	1.4	702	9	AQ391927	AQ391927	CITBI-E1-
c 316	43	1.4	704	3	BI603876	BI603876	603249768
c 317	43	1.4	705	1	AU139904	AU139904	AU139904
c 318	43	1.4	705	6	CA447891	CA447891	UI-H-EI0-
c 319	43	1.4	711	8	CX751218	CX751218	AGENCOURT
c 320	43	1.4	718	6	CD640244	CD640244	AGENCOURT
321	43	1.4	719	2	BE298219	BE298219	601118037
c 322	43	1.4	719	8	CX758818	CX758818	AGENCOURT
c 323	43	1.4	727	6	CA440514	CA440514	UI-H-ED0-
c 324	43	1.4	733	6	CD640273	CD640273	AGENCOURT
c 325	43	1.4	746	2	BE789473	BE789473	601482146
326	43	1.4	757	2	BG496744	BG496744	602538714
327	43	1.4	758	5	BX503885	BX503885	DKFZp686K
328	43	1.4	762	2	BG776280	BG776280	602663348
329	43	1.4	763	5	BU170732	BU170732	AGENCOURT
c 330	43	1.4	765	9	BZ600717	BZ600717	WHADG45TR
c 331	43	1.4	769	9	AQ491214	AQ491214	RPCI-11-2
c 332	43	1.4	771	6	CD110463	CD110463	AGENCOURT
c 333	43	1.4	779	6	CD521923	CD521923	AGENCOURT
c 334	43	1.4	815	6	CF597287	CF597287	AGENCOURT
c 335	43	1.4	838	5	BU502279	BU502279	AGENCOURT
336	43	1.4	850	8	DR763029	DR763029	HESC4_147
337	43	1.4	858	2	BE908331	BE908331	601503133
338	43	1.4	872	6	CD519236	CD519236	AGENCOURT
c 339	43	1.4	879	6	CB987766	CB987766	AGENCOURT
340	43	1.4	882	2	BE787623	BE787623	601481733
c 341	43	1.4	892	3	BQ230118	BQ230118	AGENCOURT
342	43	1.4	897	3	BM555408	BM555408	AGENCOURT
343	43	1.4	899	6	CD559035	CD559035	AGENCOURT
c 344	43	1.4	909	5	BQ881589	BQ881589	AGENCOURT
c 345	43	1.4	918	5	BX348401	BX348401	BX348401
346	43	1.4	935	2	BG432843	BG432843	602496051
347	43	1.4	940	6	CD519390	CD519390	AGENCOURT
c 348	43	1.4	966	5	BQ691900	BQ691900	AGENCOURT
349	43	1.4	970	5	BU161423	BU161423	AGENCOURT
350	43	1.4	973	5	BQ718493	BQ718493	AGENCOURT
c 351	43	1.4	987	5	BX437681	BX437681	BX437681
c 352	43	1.4	990	3	BM469408	BM469408	AGENCOURT
c 353	43	1.4	994	1	AL541554	AL541554	AL541554
354	43	1.4	1015	3	BM913110	BM913110	AGENCOURT
355	43	1.4	1023	3	BM555414	BM555414	AGENCOURT
356	43	1.4	1026	3	BM915472	BM915472	AGENCOURT
357	43	1.4	1029	3	BM471041	BM471041	AGENCOURT
358	43	1.4	1031	3	BM915471	BM915471	AGENCOURT
359	43	1.4	1044	3	BM915370	BM915370	AGENCOURT
c 360	43	1.4	1052	1	AL540730	AL540730	AL540730
c 361	43	1.4	1058	5	BX436528	BX436528	BX436528
362	43	1.4	1060	3	BM913019	BM913019	AGENCOURT
363	43	1.4	1063	3	BM807930	BM807930	AGENCOURT
364	43	1.4	1093	1	AL534375	AL534375	AL534375
c 365	43	1.4	1096	5	BX404687	BX404687	BX404687
c 366	43	1.4	1099	5	BX404833	BX404833	BX404833
367	43	1.4	1139	1	AL568568	AL568568	AL568568
368	43	1.4	1232	3	BM556405	BM556405	AGENCOURT
369	43	1.4	1294	4	BC032375	BC032375	Homo sapi
c 370	43	1.4	1701	4	CR623963	CR623963	full-leng
371	43	1.4	1706	4	CR619443	CR619443	full-leng
c 372	43	1.4	1793	4	CR620867	CR620867	full-leng
c 373	43	1.4	2263	4	AF461901	AF461901	Homo sapi
374	43	1.4	5797	4	CR749233	CR749233	Homo sapi
375	43	1.4	6345	4	HSM805983	BX537877	Homo sapi
c 376	43	1.4	7316	4	CR857086	CR857086	Pongo pyg
377	42	1.3	121	9	AQ422284	AQ422284	RPCI-11-1
378	42	1.3	144	2	BG230370	BG230370	rk34a07.y
c 379	42	1.3	145	1	AI492587	AI492587	ti29g11.x
380	42	1.3	151	6	CD642086	CD642086	AGENCOURT
381	42	1.3	183	2	BF858167	BF858167	QV1-FT020
c 382	42	1.3	230	1	AW196615	AW196615	xo25a02.x
383	42	1.3	236	8	W20344	W20344	zb44f01.r1
384	42	1.3	263	2	BE092341	BE092341	IL2-BT073
c 385	42	1.3	264	1	AW207358	AW207358	UI-H-BI1-
c 386	42	1.3	271	1	AI049508	AI049508	an21b12.x
c 387	42	1.3	300	9	AZ518434	AZ518434	RPCI-11-4

c 388	42	1.3	316	9	AQ042901	AQ042901 CIT-HSP-2
389	42	1.3	341	5	BU858574	AGENCOURT
390	42	1.3	345	1	AA569631	AA569631 nm38h07.s
391	42	1.3	349	5	BU565227	AGENCOURT
c 392	42	1.3	365	2	BE007886	QV0-BN014
c 393	42	1.3	369	2	BE007890	QV0-BN014
c 394	42	1.3	372	1	AI055937	AI055937 ox44c09.s
c 395	42	1.3	373	1	AA702614	AA702614 zf88b02.s
396	42	1.3	378	2	BF914971	BF914971 IL3-UT011
397	42	1.3	378	2	BF914982	BF914982 IL3-UT011
c 398	42	1.3	390	1	AW501938	AW501938 UI-HF-BR0
c 399	42	1.3	396	8	R02559	R02559 ye80c03.s1
400	42	1.3	404	1	AA211305	AA211305 zq86g11.r
401	42	1.3	404	3	BM991921	BM991921 UI-H-DF1-
402	42	1.3	408	1	AI088768	AI088768 qa21a07.x
c 403	42	1.3	408	1	AA099396	AA099396 zl77h06.s
c 404	42	1.3	409	1	AI627372	AI627372 ty43f09.x
c 405	42	1.3	409	9	AQ388413	AQ388413 RPCI11-15
c 406	42	1.3	411	1	AA806191	AA806191 oc25f08.s
c 407	42	1.3	419	1	AI524515	AI524515 to37f08.x
c 408	42	1.3	430	9	AQ089692	AQ089692 HS_3006_B
409	42	1.3	432	9	AQ220470	AQ220470 HS_2004_A
c 410	42	1.3	433	9	AQ017707	AQ017707 CIT-HSP-2
c 411	42	1.3	437	9	AQ774179	AQ774179 HS_2050_B
c 412	42	1.3	438	9	AQ061783	AQ061783 CIT-HSP-2
c 413	42	1.3	441	9	AQ383650	AQ383650 RPCI11-14
414	42	1.3	450	6	CD239674	CD239674 FNPBXC02
415	42	1.3	455	9	AQ359570	AQ359570 HS_5033_A
416	42	1.3	456	7	CN484836	CN484836 hx19d06.y
c 417	42	1.3	459	5	BX646449	BX646449 DKFZp781A
418	42	1.3	460	3	BQ314448	BQ314448 QV0-BN004
419	42	1.3	461	3	BQ321465	BQ321465 MR3-CT046
c 420	42	1.3	462	5	BX114476	BX114476 BX114476
c 421	42	1.3	464	1	AI536932	AI536932 tm87a01.x
c 422	42	1.3	467	9	AQ559196	AQ559196 HS_2069_B
c 423	42	1.3	468	1	AW090797	AW090797 xc96d09.x
424	42	1.3	471	9	AQ116903	AQ116903 HS_2191_A
c 425	42	1.3	479	9	B39549	B39549 HS-1049-B2-
426	42	1.3	480	9	BH140880	BH140880 UP_527-16
427	42	1.3	486	2	BE144443	BE144443 MR0-HT016
428	42	1.3	491	2	BF740930	BF740930 QV1-HB003
c 429	42	1.3	493	5	BX487887	BX487887 DKFZp686K
c 430	42	1.3	497	5	BX642370	BX642370 DKFZp686M
431	42	1.3	499	5	BX482214	BX482214 DKFZp686J
c 432	42	1.3	506	9	AQ486393	AQ486393 RPCI-11-2
433	42	1.3	507	9	AQ191426	AQ191426 HS_3233_A
434	42	1.3	513	10	CL529065	CL529065 HIV65H12.
c 435	42	1.3	515	9	AQ378058	AQ378058 RPCI-11-1
c 436	42	1.3	517	2	BE856581	BE856581 7f64h03.x
c 437	42	1.3	518	9	AQ193220	AQ193220 HS_3062_B
c 438	42	1.3	519	6	CF142069	CF142069 UI-HF-CB0
c 439	42	1.3	524	9	AQ345355	AQ345355 RPCI11-13
440	42	1.3	529	9	AQ394614	AQ394614 CITBI-E1-
c 441	42	1.3	535	9	AQ395337	AQ395337 CITBI-E1-
442	42	1.3	540	1	AU147274	AU147274 AU147274
c 443	42	1.3	542	1	AL712813	AL712813 DKFZp686E
444	42	1.3	550	7	CR545302	CR545302 DKFZp459I
c 445	42	1.3	557	9	AQ476584	AQ476584 CITBI-E1-
c 446	42	1.3	559	5	BU075290	BU075290 in29e06.y
c 447	42	1.3	561	1	AI049598	AI049598 an38a06.x
c 448	42	1.3	574	5	BX508959	BX508959 DKFZp686K
449	42	1.3	579	9	AQ052779	AQ052779 RPCI11-49
c 450	42	1.3	580	3	BP302466	BP302466 BP302466
c 451	42	1.3	580	9	AQ583761	AQ583761 RPCI-11-4
c 452	42	1.3	582	3	BP227983	BP227983 BP227983
453	42	1.3	583	9	AQ078254	AQ078254 CIT-HSP-2
c 454	42	1.3	586	9	AZ517590	AZ517590 RPCI-11-4
c 455	42	1.3	599	9	AQ548881	AQ548881 RPCI-11-3
c 456	42	1.3	600	5	BX471088	BX471088 DKFZp686F
457	42	1.3	604	9	AZ757834	AZ757834 HSC 00279
458	42	1.3	609	1	AV763305	AV763305 AV763305
c 459	42	1.3	613	9	AZ517246	AZ517246 RPCI-11-7
460	42	1.3	613	9	AQ347245	AQ347245 RPCI11-11
461	42	1.3	617	1	AI133033	AI133033 HA1621 Hu
462	42	1.3	617	7	CR753262	CR753262 DKFZp469K
463	42	1.3	620	7	CR543489	CR543489 DKFZp459M
c 464	42	1.3	621	10	AG019109	AG019109 Homo sapi
c 465	42	1.3	621	10	AG077873	AG077873 Pan trogl
c 466	42	1.3	644	9	AQ538448	AQ538448 RPCI-11-3
467	42	1.3	651	1	AV762002	AV762002 AV762002
468	42	1.3	659	10	AG142117	AG142117 Pan trogl
469	42	1.3	660	10	AG045128	AG045128 Pan trogl
c 470	42	1.3	664	5	BX471890	BX471890 DKFZp686A
c 471	42	1.3	665	7	CN304726	CN304726 170005977
c 472	42	1.3	665	10	AG079274	AG079274 Pan trogl
473	42	1.3	668	5	BX506469	BX506469 DKFZp779E
c 474	42	1.3	669	10	AG019043	AG019043 Homo sapi
c 475	42	1.3	678	5	BX376229	BX376229 BX376229
476	42	1.3	684	10	AG045683	AG045683 Pan trogl
477	42	1.3	688	6	CD516136	CD516136 AGENCOURT
c 478	42	1.3	689	1	AL704286	AL704286 DKFZp686P
479	42	1.3	693	10	AG176532	AG176532 Pan trogl
480	42	1.3	695	5	BX376230	BX376230 BX376230
481	42	1.3	696	4	CR616459	CR616459 full-leng
c 482	42	1.3	700	6	CA777576	CA777576 ip18c02.y
483	42	1.3	705	1	AV760497	AV760497 AV760497
c 484	42	1.3	707	9	AQ262047	AQ262047 CITBI-E1-
485	42	1.3	718	7	CR745334	CR745334 CR745334
c 486	42	1.3	722	7	CR856700	CR856700 DKFZp469H
487	42	1.3	727	10	AG187192	AG187192 Pan trogl
488	42	1.3	728	1	AV763952	AV763952 AV763952
c 489	42	1.3	728	2	BE889588	BE889588 601512509
c 490	42	1.3	728	9	AQ554491	AQ554491 RPCI-11-4
491	42	1.3	740	6	CB986442	CB986442 AGENCOURT
492	42	1.3	741	11	CR958354	CR958354 Homo sapi
c 493	42	1.3	746	7	CN278648	CN278648 170006003
c 494	42	1.3	751	6	CB998596	CB998596 AGENCOURT
c 495	42	1.3	751	9	B60698	B60698 CIT-HSP-201
c 496	42	1.3	757	1	AU122466	AU122466 AU122466
497	42	1.3	761	7	CR559036	CR559036 DKFZp468P
498	42	1.3	781	10	CG846669	CG846669 HSC 00933
499	42	1.3	784	6	CB962505	CB962505 AGENCOURT
c 500	42	1.3	797	2	BG778173	BG778173 602666325

ALIGNMENTS

RESULT 1
BX116028
LOCUS BX116028 660 bp mRNA linear EST 07-FEB-2003
DEFINITION BX116028 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGp998C11795 ; IMAGE:347458, mRNA sequence.
ACCESSION BX116028
VERSION BX116028.1 GI:27880354
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998C11795.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany

```
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .660
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGp998C11795 ; IMAGE:347458"
                      /sex="unknown"
                      /dev_stage="19 weeks"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares_fetal_heart_NbHH19W"
                      /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by
                      M.Fatima Bonaldo. This library was constructed from the
                      same fetus as the fetal lung library, Soares fetal lung
                      NbHL19W."
ORIGIN

  Query Match            20.7%;  Score 646;  DB 5;  Length 660;
  Best Local Similarity  100.0%;  Pred. No. 8e-303;
  Matches  646;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2129 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 60

Qy       2189 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 2248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 120

Qy       2249 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 180

Qy       2309 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 2368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 240

Qy       2369 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 2428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAG 300

Qy       2429 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 2488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCAC 360

Qy       2489 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 2548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAG 420

Qy       2549 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 2608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAG 480

Qy       2609 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 2668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTA 540

Qy       2669 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 2728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACT 600

Qy       2729 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             |||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 646


RESULT 2
CF145408
LOCUS       CF145408                670 bp    mRNA    linear   EST 06-AUG-2003
DEFINITION  UI-HF-CB0-aso-e-10-0-UI.r1 NIH_MGC_210 Homo sapiens cDNA clone
            IMAGE:30569145 5', mRNA sequence.
ACCESSION   CF145408
VERSION     CF145408.1  GI:33260852
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 670)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Tim Ratlift
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDNA
            sequence: 129-236, >MIR#SINE/MIR
            Seq primer: pYX-5.
FEATURES              Location/Qualifiers
     source           1. .670
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30569145"
                      /tissue_type="CNCAP(3)T-225 cell line"
                      /lab_host="DH10B (T1 phage resistant)"
                      /clone_lib="NIH_MGC_210"
                      /note="Organ: Prostate; Vector: pT7T3 Pac; Site_1: EcoR I;
                      Site_2: Not I; The library was constructed according
                      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                      1996. Denatured RNA was size fractionated on a 1% agarose
                      gel. First strand cDNA synthesis was primed with oligo-dT
                      primer containing a Not I site. Double strand cDNA was
                      size selected according to mRNA size fraction, ligated
                      with EcoR I adaptor, digested with Not I and then cloned
                      directionally into pT7T3 Pac vector. The library tag
                      sequence located between the Not I site and the polyA tail
                      is CCCAC. Tissue was provided by Tim Ratlift."
ORIGIN

  Query Match            19.6%;  Score 613;  DB 6;  Length 670;
  Best Local Similarity  99.8%;  Pred. No. 9.6e-287;
  Matches  663;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         43 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 66

Qy        103 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         67 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 126

Qy        163 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 186

Qy        223 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 246

Qy        283 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 306

Qy        343 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 366

Qy        403 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 426

Qy        463 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 486
.
Qy        523 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 546

Qy        583 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTA 642
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        547 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGATGGAGCGGGGGACCGCTA 606

Qy        643 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCT 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCT 666

Qy        703 CCGT 706
              ||||
Db        667 CCGT 670


RESULT 3
CF145448
LOCUS       CF145448                677 bp    mRNA    linear   EST 06-AUG-2003
DEFINITION  UI-HF-CB0-asp-a-08-0-UI.r1 NIH_MGC_210 Homo sapiens cDNA clone
            IMAGE:30569191 5', mRNA sequence.
ACCESSION   CF145448
VERSION     CF145448.1  GI:33260892
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Tim Ratlift
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDNA
            sequence: 129-236, >MIR#SINE/MIR
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30569191"
                     /tissue_type="CNCAP(3)T-225 cell line"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_210"
                     /note="Organ: Prostate; Vector: pT7T3 Pac; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pT7T3 Pac vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CCCAC. Tissue was provided by Tim Ratlift."
ORIGIN

  Query Match             19.2%;  Score 598;  DB 6;  Length 677;
  Best Local Similarity   99.8%;  Pred. No. 2e-279;
  Matches  648;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         43 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 GCGCAACGCGGCAGGGACACCTGACCCCGGCGGCGCCCAGCCCCTCGGATTGCCAGTCAC 66

Qy        103 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 TGCTCGCTTTGGGGCACGGAGGTGCCCAGTCCTGCGGGGCACCCGACGTCCTGTCGCCGA 126

Qy        163 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGGGTCCGGGAGTCAGTATAGCTGGGTTCTAGTCCCATCACAGGCAAAAACTCCGCGGG 186

Qy        223 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCCGTAAAATAGAACGGGTTGG 246

Qy        283 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAAGGGGAGGGAGGGATGGGCC 306

Qy        343 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 ACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACTGATCCAGGGGGTGGCAGCT 366

Qy        403 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAG 426

Qy        463 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAG 486

Qy        523 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAG 546

Qy        583 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTA 642
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        547 ACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGAACGCTA 606

Qy        643 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTC 691
              |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        607 ACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTC 655

RESULT 4
BM671616/c
LOCUS       BM671616                685 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CQ1-agc-l-04-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
            UI-E-CQ1-agc-l-04-0-UI 3', mRNA sequence.
ACCESSION   BM671616
VERSION     BM671616.1  GI:18981514
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CQ1-agc-l-04-0-UI"
                     /tissue_type="optic nerve"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CQ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CQ1 is a normalized cDNA library containing the
                     following tissue(s): optic nerve. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     CCATTAAGTG. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI).
                     TAG_TISSUE=human optic nerve
                     TAG_LIB=UI-E-CQ1
                     TAG_SEQ=CCATTAAGTG"
ORIGIN

  Query Match           17.0%;  Score 531;  DB 3;  Length 685;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-247;
  Matches  651;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       2121 GGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGA 2180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 GGCTTTCTCTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGA 609

Qy       2181 GCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTT 2240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GCAGCTAGAAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTT 549

Qy       2241 AGGAAACCAAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACT 2300
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        548 AGGAAACCAAAGAGGGTGGG-TGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACT 490

Qy       2301 TGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAG 2360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 TGCCTGCGGAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAG 430

Qy       2361 GCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTT 2420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GCTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTT 370

Qy       2421 TTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTT 2480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TTGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTT 310

Qy       2481 TAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGC 2540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 TAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGC 250

Qy       2541 TTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGC 2600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TTTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGC 190

Qy       2601 CCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTT 2660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CCACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTT 130

Qy       2661 GGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAG 2720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GGGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAG 70

Qy       2721 CGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCA 2772
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          69 CGTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCA 18

RESULT 5
BM707056
LOCUS       BM707056                578 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-CR1-adx-b-12-0-UI.r1 UI-E-CR1 Homo sapiens cDNA clone
            UI-E-CR1-adx-b-12-0-UI 5', mRNA sequence.
ACCESSION   BM707056
VERSION     BM707056.1  GI:19020314
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
```

```
            DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
            Clone Distribution: Researchers may obtain clones from Research
           Genetics (www.resgen.com).
           Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CR1-adx-b-12-0-UI"
                     /tissue_type="eye anterior segment"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CR1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CR1 is a normalized cDNA library containing the
                     following tissue(s): eye anterior segment. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AATGCCGCAT. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI)."
ORIGIN

  Query Match             16.6%;  Score 519;  DB 3;  Length 578;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-241;
  Matches  519;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1964 GGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCG 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTTTCCCG 77

Qy       2024 GACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCCTA 137

Qy       2084 AATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACA 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACA 197

Qy       2144 GCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTC 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGGTC 257

Qy       2204 ATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTG 2263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGTTG 317

Qy       2264 AACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCC 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTCCC 377

Qy       2324 CAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTG 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGGTG 437

Qy       2384 TTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTG 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAGTG 497

Qy       2444 TTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTA 2482
             |||||||||||||||||||||||||||||||||||||||
Db        498 TTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTA 536
```

```
RESULT 6
BM710194
LOCUS       BM710194                500 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-CQ1-agc-l-04-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone
            UI-E-CQ1-agc-l-04-0-UI 5', mRNA sequence.
ACCESSION   BM710194
VERSION     BM710194.1  GI:19023452
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .500
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CQ1-agc-l-04-0-UI"
                     /tissue_type="optic nerve"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CQ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CQ1 is a normalized cDNA library containing the
                     following tissue(s): optic nerve. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     CCATTAAGTG. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI)."
ORIGIN

  Query Match             16.0%;  Score 500;  DB 3;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-231;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2022 CGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCC 2081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGGGCC 60

Qy       2082 TAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTA 2141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTA 120
```

```
Qy       2142 CAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGG 2201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTGAGG 180

Qy       2202 TCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGT 2261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTGGGT 240

Qy       2262 TGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTC 2321
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCGGAAAAGGGCTCTC 300

Qy       2322 CCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGG 2381
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGTAGGGCCACCTGG 360

Qy       2382 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 2441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGAACTAGACCAGAG 420

Qy       2442 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 2501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTA 480

Qy       2502 TCCCCTGGAATGTGCGTGCT 2521
             ||||||||||||||||||||
Db        481 TCCCCTGGAATGTGCGTGCT 500
```

RESULT 7
BQ187216

```
LOCUS       BQ187216                655 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-E-EJ1-ajz-e-14-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajz-e-14-0-UI 5', mRNA sequence.
ACCESSION   BQ187216
VERSION     BQ187216.1  GI:20362767
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 296-333, >GC_rich#Low_complexity (matched compliment)
            Seq primer: M13 REVERSE.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajz-e-14-0-UI"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-EJ1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes,
                     AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
                     AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
                     Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
                     library was created for the program, Gene Discovery in the
                     Visual System, supported by National Eye Institute (NEI)."
ORIGIN
```

```
  Query Match             14.4%;  Score 450;  DB 3;  Length 655;
  Best Local Similarity   99.4%;  Pred. No. 2.6e-207;
  Matches  650;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        526 GGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACA 585
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          2 GGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCANAGACA 61

Qy        586 CGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACG 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACG 121

Qy        646 GGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCG 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCG 181

Qy        706 TGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGG 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGG 241

Qy        766 CCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCG 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCG 301

Qy        826 CGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGC 885
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |
Db        302 CGGGCGGCCGGGCCCAGCCGGAGCCCACCGCGATGGCNAGGGAGGAGTGCAAGGCGCTTC 361

Qy        886 TGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCT 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGGACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCT 421

Qy        946 CGGCGGACTCGCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGC 1005
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CGGCGGACTCGCANAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGC 481

Qy       1006 TGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCG 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TGGCGGTGTCCACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCG 541

Qy       1066 ACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGG 601

Qy       1126 AAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGC 1179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AAGCGGACATGCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGC 655
```

RESULT 8
AW302149/c

```
LOCUS       AW302149                476 bp    mRNA    linear   EST 18-JAN-2000
DEFINITION  xs01f06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2768387 3'
            similar to contains TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AW302149
VERSION     AW302149.1  GI:6711826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .476
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2768387"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
ORIGIN

  Query Match             13.7%;  Score 427;  DB 1;  Length 476;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-196;
  Matches  427;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 417

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 357

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    .   356 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 297

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 237

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 177

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 117

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 57

Qy       1136 GCGACGC 1142
              |||||||
Db          56 GCGACGC 50


RESULT 9
AI375213/c
LOCUS       AI375213                455 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  tc10f06.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2063459
            3', mRNA sequence.
ACCESSION   AI375213
VERSION     AI375213.1  GI:4175203
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 587   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2063459"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
ORIGIN

  Query Match             13.2%;  Score 413;  DB 1;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-189;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2362 CTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTT 2421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CTTGGGGTAGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTT 356

Qy       2422 TGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTT 2481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TGTTCAGAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTT 296
```

```
Qy       2482 AGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCT 2541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 AGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCT 236

Qy       2542 TTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCC 2601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TTGGCAGCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCC 176

Qy       2602 CACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTG 2661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CACAGAGGTTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTG 116

Qy       2662 GGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGC 2721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 GGATTTAAGAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGC 56

Qy       2722 GTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          55 GTGTACTATGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCATA 3
```

```
RESULT 10
BM673230/c
LOCUS       BM673230                480 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CR1-adx-b-12-0-UI.s1 UI-E-CR1 Homo sapiens cDNA clone
            UI-E-CR1-adx-b-12-0-UI 3', mRNA sequence.
ACCESSION   BM673230
VERSION     BM673230.1  GI:18983128
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CR1-adx-b-12-0-UI"
                     /tissue_type="eye anterior segment"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CR1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CR1 is a normalized cDNA library containing the
                     following tissue(s): eye anterior segment. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AATGCCGCAT. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI).
                     TAG_TISSUE=human eye anterior segment
                     TAG_LIB=UI-E-CR1
                     TAG_SEQ=AATGCCGCAT"
ORIGIN
```

```
  Query Match             13.2%;  Score 412;  DB 3;  Length 480;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-189;
  Matches  462;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2310 AAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGT 2369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGGT 421

Qy       2370 AGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGA 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AGGGCCACCTGGTGTTTAAACAGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCAGA 361

Qy       2430 ACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACT 2489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCACT 301

Qy       2490 TAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGC 2549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCAGC 241

Qy       2550 TCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGG 2609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TCCCTGACCCCCGCGCTGCCCGCCCCTCCGGGGTAATGTGGCATTACTGGCCCACAGAGG 181

Qy       2610 TTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAA 2669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TTTTGAGCCAATCAGCTCTGAGACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAA 121

Qy       2670 GAAGCTTTTAAAAGGTAATAATCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTA 2729
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        120 GAAGCTTTTAAAAGGTAATAATACTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTA 61

Qy       2730 TGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCA 2772
             |||||||||||||||||||||||||||||||||||||||||||
Db          60 TGAAAGCTGTTATTTTAATAAAGAACGCTGGGCCATGAACTCA 18
```

```
RESULT 11
BQ682843
LOCUS       BQ682843                972 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8495014 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6295149
            5', mRNA sequence.
ACCESSION   BQ682843
VERSION     BQ682843.1  GI:21795522
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 972)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM2500  row: o  column: 22
           High quality sequence stop: 463.
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6295149"
                     /tissue_type="melanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_112"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             12.7%;  Score 395;  DB 5;  Length 972;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-180;
  Matches  395;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 AGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGT 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGTTCGGTGCCCACCCCTGCCAGCGTCGCCCCCTTTCTGCGTGGGACAGTTTGAAAAGGT 60

Qy       1959 GGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTT 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGTGGGGTGGAGTGAAGTTTGGAGAGGGACGCTGTTTGGTTCTATGTGGTTGGTCTGTT 120

Qy       2019 TCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCCGGACAAGAAAAATTGCAATCAAATGTCAGCAGCTTTTATTACCTTAATCTTTCAGG 180

Qy       2079 GCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTAAATTTAGGAGAGTGTCCTGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCG 240

Qy       2139 CTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTG 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAGAAGAGATTTG 300

Qy       2199 AGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTG 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACCAAAGAGGGTG 360

Qy       2259 GGTTGAACCTACTCTCACGGACTTGGATCCAGTGC 2293
             |||||||||||||||||||||||||||||||||||
Db        361 GGTTGAACCTACTCTCACGGACTTGGATCCAGTGC 395


RESULT 12
BM562937
LOCUS       BM562937                1376 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6566935 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5736787
            5', mRNA sequence.
ACCESSION   BM562937
VERSION     BM562937.1  GI:18809436
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1376)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM12745  row: n  column: 20
           High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .1376
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5736787"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             11.1%;  Score 348;  DB 3;  Length 1376;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-157;
  Matches  348;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        453 CCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGC 512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGC 60

Qy        513 ACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGG 572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGGAGGCCAGGG 120

Qy        573 CGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGG 632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACGAGGGAGCGG 180

Qy        633 GGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCC 692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGTCGAGGGTCC 240

Qy        693 CGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGT 752
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGT 300

Qy        753 TCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCC 800
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCGTGCCGCGCGGCCCAAGGCCGGGATGGGGGTTAGCCACATCCTGCC 348


RESULT 13
W81663
LOCUS       W81663                  595 bp    mRNA    linear   EST 17-OCT-1996
DEFINITION  zd85e06.r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:347458 5', mRNA sequence.
ACCESSION   W81663
VERSION     W81663.1  GI:1392291
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
```

```
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 584   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1272833"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:347458"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
ORIGIN

  Query Match             10.7%;  Score 333;  DB 8;  Length 595;
  Best Local Similarity   99.8%;  Pred. No. 2.7e-150;
  Matches  453;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       2129 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGACGCGCTACAGCCCTTCCTAGCAGAGTTTATCCATTCGTCCCCAAGAGCAGCTAG 60

Qy       2189 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 2248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCTATTTAGGAAACC 120

Qy       2249 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 2308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGGGTGGGTTGAACCTACTCTCACGGACTTGGATCCAGTGCGCACACTTGCCTGCG 180

Qy       2309 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 2368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAAGGGCTCTCCCCAGCCACCCGGAGATGGGGGTAAGAGGAAGAGCAGAGGCTTGGGG 240

Qy       2369 TAGGGCCACCTGGTGTTTAAACA-GGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCA 2427
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        241 TAGGGCCACCTGGTGTTTAAACACGGCACTTTCTCCTTCTCTGGGGCTTATTTTTGTTCA 300

Qy       2428 GAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCA 2487
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACTAGACCAGAGTGTTTGAACCTCCTTTGCAGGAGGGCTGGGAATCCTCTTTAGAGCA 360

Qy       2488 CTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCA 2547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGGCCAGTAGGAGGGCTGGCTTTGGCA 420

Qy       2548 GCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGG 2581
             |||||||||||||||||||||||||||||||||
Db        421 GCTCCCTGACCCCCGCGCTGCCCGCCCCTCCGGG 454


RESULT 14
BI458731/c
LOCUS       BI458731                605 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603199879F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5265979 5',
            mRNA sequence.
ACCESSION   BI458731
VERSION     BI458731.1  GI:15249387
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11670  row: m  column: 20
            High quality sequence stop: 591.
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5265979"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

  Query Match             9.3%;  Score 290;  DB 3;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-129;
  Matches  290;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        422 TGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGC 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TGGGCGGGTCCTAGGAAACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGC 241

Qy        482 GGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCA 181

Qy        542 GCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGA 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GCTTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGA 121

Qy        602 GTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCG 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        120 GTGCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCG 61

Qy        662 CCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACG 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 CCCCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACG 11


RESULT 15
BX645932
LOCUS       BX645932                728 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp781B1952_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781B1952 5', mRNA sequence.
ACCESSION   BX645932
VERSION     BX645932.1  GI:34480265
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781B1952) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781B1952"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             5.8%;  Score 180;  DB 5;  Length 728;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-76;
  Matches  180;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2693 CTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAG 2752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 CTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAAAG 80

Qy       2753 AACGCTGGGCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATA 2812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AACGCTGGGCCATGAACTCATACCTGCCAATGAGTCAAACATAGTATCTTTATGTAGATA 140

Qy       2813 CTTAGATTACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCCC 2872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 CTTAGATTACTAAATATATATTTCATCTACTTCTGAAGTTGATAGTCTTCCCCCCCCCCC 200


RESULT 16
W81222/c
LOCUS       W81222                  449 bp    mRNA    linear   EST 17-OCT-1996
DEFINITION  zd85e06.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:347458 3', mRNA sequence.
ACCESSION   W81222
VERSION     W81222.1  GI:1392242
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 584   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 76.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:1272833"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:347458"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
ORIGIN

  Query Match             4.6%;  Score 144;  DB 8;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-58;
  Matches  144;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2631 GACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAA 2690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GACTGGGTTAGAATGTAACAGCTTTAACTTGGGATTTAAGAAGCTTTTAAAAGGTAATAA 86

Qy       2691 TCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAA 2750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 TCCTCTGAAAGAAAAATGACGTAACCACAGCGTGTACTATGAAAGCTGTTATTTTAATAA 26

Qy       2751 AGAACGCTGGGCCATGAACTCATA 2774
              ||||||||||||||||||||||||
Db         25 AGAACGCTGGGCCATGAACTCATA 2


RESULT 17
AG013685
```

```
LOCUS       AG013685                709 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
            survey sequence.
ACCESSION   AG013685
VERSION     AG013685.2  GI:55788901
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 709)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
COMMENT     On Nov 16, 2004 this sequence version replaced gi:3551633.
FEATURES             Location/Qualifiers
     source          1. .709
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="B137B7N13"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 10;  Length 709;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-14;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 445


RESULT 18
AG013687
LOCUS       AG013687                709 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
            survey sequence.
ACCESSION   AG013687
VERSION     AG013687.1  GI:3551635
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 709)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
FEATURES             Location/Qualifiers
     source          1. .709
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="B137B7N13"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 10;  Length 709;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-14;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 388


RESULT 19
AG013686
LOCUS       AG013686                718 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
            survey sequence.
ACCESSION   AG013686
VERSION     AG013686.1  GI:3551634
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 718)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="B137B7N13"
ORIGIN

  Query Match             1.7%;  Score 53;  DB 10;  Length 718;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-14;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 448


RESULT 20
AA679936/c
LOCUS       AA679936                354 bp    mRNA    linear   EST 02-DEC-1997
DEFINITION  ac85g01.s1 Stratagene ovary (#937217) Homo sapiens cDNA clone
            IMAGE:869424 3' similar to contains Alu repetitive element;contains
            element TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AA679936
VERSION     AA679936.1  GI:2656403
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
```

```
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 285.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:869424"
                     /sex="female"
                     /dev_stage="49 year old"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene ovary (#937217)"
                     /note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
                     Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Total ovary tissue, normal, caucasian. Average insert
                     size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 1;  Length 354;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 166


RESULT 21
AI732120/c
LOCUS       AI732120                441 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  ac86a01.x5 Stratagene ovary (#937217) Homo sapiens cDNA clone
            IMAGE:869448 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AI732120
VERSION     AI732120.1  GI:5053255
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced human clone
            Original clone citation: WashU-NCI human EST Project
            This read has been verified (found to hit its original self in the
            correct orientation)
            Insert Length: 1252   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 437.
FEATURES             Location/Qualifiers
     source          1. .441
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:869448"
                     /sex="female"
                     /dev_stage="49 year old"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene ovary (#937217)"
                     /note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
                     Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Total ovary tissue, normal, caucasian. Average insert
                     size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 1;  Length 441;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 166


RESULT 22
AI732180/c
LOCUS       AI732180                441 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  ac85g01.x5 Stratagene ovary (#937217) Homo sapiens cDNA clone
            IMAGE:869424 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AI732180
VERSION     AI732180.1  GI:5053293
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced human clone
            Original clone citation: WashU-NCI human EST Project
            This read has been verified (found to hit its original self in the
            correct orientation)
            Insert Length: 1299   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .441
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:869424"
                     /sex="female"
                     /dev_stage="49 year old"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene ovary (#937217)"
                     /note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
                     Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Total ovary tissue, normal, caucasian. Average insert
                     size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
```

```
ORIGIN

  Query Match             1.7%;  Score 52;  DB 1;  Length 441;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 166


RESULT 23
CN277748
LOCUS       CN277748                 520 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600183571 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN277748
VERSION     CN277748.1  GI:47294162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 520   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 7;  Length 520;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 518


RESULT 24
AA680243/c
LOCUS       AA680243                 529 bp    mRNA    linear   EST 02-DEC-1997
DEFINITION  ac86a01.s1 Stratagene ovary (#937217) Homo sapiens cDNA clone
            IMAGE:869448 3' similar to contains Alu repetitive element;contains
            element L1 repetitive element ;, mRNA sequence.
ACCESSION   AA680243
VERSION     AA680243.1  GI:2656211
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:869448"
                     /sex="female"
                     /dev_stage="49 year old"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Stratagene ovary (#937217)"
                     /note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
                     Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Total ovary tissue, normal, caucasian. Average insert
                     size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 192


RESULT 25
CK825142/c
LOCUS       CK825142                 553 bp    mRNA    linear   EST 11-MAR-2004
DEFINITION  ig44h08.y5 HR85 islet Homo sapiens cDNA clone IMAGE:5594343 5',
            mRNA sequence.
ACCESSION   CK825142
VERSION     CK825142.1  GI:44842067
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
            Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
            Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
            McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
            Bowers,Y.
  TITLE     WashU-Harvard Pancreas EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
```

```
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            This read is a 5' RESEQUENCE of a previously sequenced pancreas
            clone
            Good hit to opposite strand read. . .wrong orientation BUT PASSED FOR
            MOUSE-PANCREAS VERIFICATION
            Seq primer: -40UP from Gibco
            High quality sequence stop: 553.
FEATURES              Location/Qualifiers
     source           1. .553
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5594343"
                      /tissue_type="Purified pancreatic islet"
                      /lab_host="DH10B"
                      /clone_lib="HR85 islet"
                      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                      Size-selected on agarose gel. Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                      314-362-1916, Fax: 314-747-2692."
ORIGIN

  Query Match             1.7%;  Score 52;  DB 7;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 489


RESULT 26
CK825141
LOCUS       CK825141                555 bp    mRNA    linear   EST 11-MAR-2004
DEFINITION  ig44h08.x5 HR85 islet Homo sapiens cDNA clone IMAGE:5594343 3',
            mRNA sequence.
ACCESSION   CK825141
VERSION     CK825141.1  GI:44842066
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
            Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
            Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
            McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
            Bowers,Y.
  TITLE     WashU-Harvard Pancreas EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig44h08.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            This read is a 3' RESEQUENCE of a previously sequenced pancreas
            clone
            Good hit to opposite strand read. . .wrong orientation BUT PASSED FOR
            MOUSE-PANCREAS VERIFICATION
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .555
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5594343"
                      /tissue_type="Purified pancreatic islet"
                      /lab_host="DH10B"
                      /clone_lib="HR85 islet"
                      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                      Size-selected on agarose gel. Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                      314-362-1916, Fax: 314-747-2692."
ORIGIN

  Query Match             1.7%;  Score 52;  DB 7;  Length 555;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 491


RESULT 27
AQ526715
LOCUS       AQ526715                568 bp    DNA     linear   GSS 11-MAY-1999
DEFINITION  HS_5216_A2_C08_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=792 Col=16 Row=E, genomic survey sequence.
ACCESSION   AQ526715
VERSION     AQ526715.1  GI:4769865
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 792  row: E  column: 16
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 568.
FEATURES              Location/Qualifiers
     source           1. .568
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=792 Col=16 Row=E"
                      /sex="male"
```

```
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.7%;  Score 52;  DB 9;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-13;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 205


RESULT 28
BZ606000/c
LOCUS       BZ606000                373 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHAAS80TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-5N15, genomic survey sequence.
ACCESSION   BZ606000
VERSION     BZ606000.1  GI:31514561
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 373)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .373
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-5N15"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN

  Query Match             1.6%;  Score 51;  DB 9;  Length 373;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 135


RESULT 29
CR545224
LOCUS       CR545224                350 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  DKFZp459E1025_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone
            DKFZp459E1025 5', mRNA sequence.
ACCESSION   CR545224
VERSION     CR545224.1  GI:49897049
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Koehrer,K., Beyer,A., Mewes,H.W., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical
            Research Center at the Heinrich-Heine-University,
            Duesseldorf/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp459E1025) is available at
            the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
            Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
            clone@rzpd.de Further information about the clone and the
            sequencing project is available at
            http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .350
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp459E1025"
                     /tissue_type="cortex"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="459 (synonym: pcor1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             1.6%;  Score 50;  DB 7;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 147


RESULT 30
CB140816
LOCUS       CB140816                471 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0194308 L15CKK1 Homo sapiens cDNA clone L15CKK1-33-C11 5',
            mRNA sequence.
ACCESSION   CB140816
VERSION     CB140816.1  GI:28115050
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
```

```
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 33  row: C  column: 11
            High quality sequence stop: 471.
FEATURES              Location/Qualifiers
     source           1. .471
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="L15CKK1-33-C11"
                      /sex="M"
                      /cell_line="CK-K1"
                      /lab_host="Top10F'"
                      /clone_lib="L15CKK1"
                      /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                      Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                      bacterial alkaline phosphatase (BAP) and then decapped
                      with tabacco acid pyrophosphatase (TAP). The decapped
                      intact mRNA was ligated with DNA-RNA linker including
                      EcoRI site by treatment of T4 RNA ligase and the first
                      strand cDNA was synthesized from oligo dT-selected mRNA by
                      priming with dT-tailed vector. The dT-tailed vector was
                      adjusted to have about 60nt. The cDNA vector was
                      circularized with E. coli DNA ligase after digestion of
                      EcoRI which site is also included in vector. An RNA strand
                      converted to a DNA strand by Okayama-Berg method. The
                      obtained cDNA vectors were used for transformation of
                      competent cells E. coli Top10F' by electroporation method.
                      The cDNA libraries constructed by this method are
                      full-length enriched cDNA library."
ORIGIN

  Query Match             1.6%;  Score 50;  DB 6;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2973 AATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 3022
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 AATTGGCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 194


RESULT 31
AQ487295/c
LOCUS       AQ487295                572 bp    DNA     linear   GSS 24-APR-1999
DEFINITION  RPCI-11-245E13.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-245E13, genomic survey sequence.
ACCESSION   AQ487295
VERSION     AQ487295.1  GI:4673169
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-245E13.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .572
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7593804"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-245E13"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /clone_lib="RPCI-11"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.6%;  Score 50;  DB 9;  Length 572;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 523


RESULT 32
B47416/c
LOCUS       B47416                  634 bp    DNA     linear   GSS 18-JUN-1998
DEFINITION  CIT978SK-161J21.TR CIT978SK Homo sapiens genomic clone 161J21,
            genomic survey sequence.
ACCESSION   B47416
VERSION     B47416.1  GI:2599822
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT978SK-161J21.TVB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .634
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:5283048"
                      /db_xref="taxon:9606"
                      /clone="161J21"
                      /sex="Female"
                      /cell_type="Fibroblast"
                      /clone_lib="CIT978SK"
                      /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
```

```
                    CalTech Human BAC Library A"
ORIGIN

  Query Match             1.6%;  Score 50;  DB 9;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 537


RESULT 33
AG145253
LOCUS       AG145253                656 bp    DNA     linear   GSS 08-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-006H23.T7, genomic survey
            sequence.
ACCESSION   AG145253
VERSION     AG145253.1  GI:16674931
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                Sequencing: T7
             LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-006H23.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.6%;  Score 50;  DB 10;  Length 656;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 443


RESULT 34
AV734315
LOCUS       AV734315                694 bp    mRNA    linear   EST 17-OCT-2000
DEFINITION  AV734315 cdA Homo sapiens cDNA clone cdAANC05 5', mRNA sequence.
ACCESSION   AV734315
VERSION     AV734315.1  GI:10851860
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S.,
            Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
            Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA cdA clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cdAANC05"
                     /tissue_type="pheochromocytoma"
                     /dev_stage="Adult"
                     /lab_host="BM25.8"
                     /clone_lib="cdA"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

  Query Match             1.6%;  Score 50;  DB 1;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 446


RESULT 35
CN413517
LOCUS       CN413517                706 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000532630078 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN413517
VERSION     CN413517.1  GI:47401111
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 706)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 706   Std Error: 0.00.
FEATURES             Location/Qualifiers
```

```
     source          1. .706
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

  Query Match             1.6%;  Score 50;  DB 7;  Length 706;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
-             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 183


RESULT 36
CD357076
LOCUS       CD357076                770 bp    mRNA    linear   EST 29-MAY-2003
DEFINITION  AGENCOURT_14249993 NIH_MGC_187 Homo sapiens cDNA clone
            IMAGE:30401040 5', mRNA sequence.
ACCESSION   CD357076
VERSION     CD357076.1  GI:31128487
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM181  row: c  column: 01
            High quality sequence stop: 538.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30401040"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_187"
                     /note="Organ: blood vessels - aorta, basilar and artery;
                     Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc); Site_2:
                     SfiI (ggccgcctcggcc); 5' and 3' adaptors were used in
                     cloning as follows: 5' adaptor sequence:
                     5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.4 kb
                     (range 0.5-4.0 kb). 14/15 colonies contained inserts by
                     PCR. This library was enriched for full-length clones and
                     was constructed by Clontech Laboratories (Palo Alto, CA).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             1.6%;  Score 50;  DB 6;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 431


RESULT 37
AY758761
LOCUS       AY758761                797 bp    DNA     linear   GSS 14-APR-2005
DEFINITION  RP43-99b9 TV Pan troglodytes genomic DNA Pan troglodytes genomic
            clone RP43-99b9 TV, genomic survey sequence.
ACCESSION   AY758761
VERSION     AY758761.1  GI:62544847
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Yohn,C.T., Jiang,Z., McGrath,S.D., Hayden,K.E., Khaitovich,P.,
            Johnson,M.E., Eichler,M.Y., Mcpherson,J.D., Zhao,S., Paabo,S. and
            Eichler,E.E.
  TITLE     Lineage-specific expansions of retroviral insertions within the
            genomes of African great apes but not humans and orangutans
  JOURNAL   PLoS Biol. 3 (4), e110 (2005)
   PUBMED   15737067
COMMENT     Contact: Evan Eichler
            Department of Genetics
            Case Western Reserve University
            2109 Adelbert Rd., BRB 720, Cleveland, OH 44106, USA
            Tel: 216 368 4883
            Fax: 216 368 3432
            Email: eee@po.cwru.edu
            TJ forward and TV reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-99b9 TV"
                     /clone_lib="Pan troglodytes genomic DNA"
ORIGIN

  Query Match             1.6%;  Score 49;  DB 9;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-12;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 452


RESULT 38
AW006453/c
LOCUS       AW006453                452 bp    mRNA    linear   EST 10-SEP-1999
DEFINITION  wr16h01.x1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:2481745 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AW006453
VERSION     AW006453.1  GI:5855231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 423.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2481745"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr22"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.5%;  Score 48;  DB 1;  Length 452;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-11;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 26


RESULT 39
AQ757333/c
LOCUS       AQ757333                548 bp    DNA     linear   GSS 27-JUL-1999
DEFINITION  HS_5484_A1_H10_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1060 Col=19 Row=O, genomic survey sequence.
ACCESSION   AQ757333
VERSION     AQ757333.1  GI:5622184
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 548)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1060  row: O  column: 19
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 548.
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=1060 Col=19 Row=O"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.5%;  Score 48;  DB 9;  Length 548;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-11;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 399


RESULT 40
AQ121008/c
LOCUS       AQ121008                457 bp    DNA     linear   GSS 22-SEP-1998
DEFINITION  HS_3073_A2_C01_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3073 Col=2 Row=E, genomic survey
            sequence.
ACCESSION   AQ121008
VERSION     AQ121008.1  GI:3498174
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3073  row: E  column: 2
            Class: BAC ends
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3073 Col=2 Row=E"
                     /sex="male"
```

```
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.5%;  Score 47;  DB 9;  Length 457;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 151


RESULT 41
AQ029209/c
LOCUS       AQ029209                470 bp    DNA     linear   GSS 14-APR-1999
DEFINITION  RPCI11-41A14.TP RPCI-11 Homo sapiens genomic clone RPCI-11-41A14,
            genomic survey sequence.
ACCESSION   AQ029209
VERSION     AQ029209.1  GI:3274340
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7515373"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-41A14"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 47;  DB 9;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-11;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 104


RESULT 42
AQ508783/c
LOCUS       AQ508783                566 bp    DNA     linear   GSS 29-APR-1999
DEFINITION  RPCI-11-280C7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-280C7,
            genomic survey sequence.
ACCESSION   AQ508783
VERSION     AQ508783.1  GI:4713530
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-280C7.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7607190"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-280C7"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 47;  DB 9;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-11;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 119


RESULT 43
BF526343/c
LOCUS       BF526343                847 bp    mRNA    linear   EST 11-DEC-2000
DEFINITION  602070608F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4213848
            5', mRNA sequence.
ACCESSION   BF526343
VERSION     BF526343.1  GI:11613706
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9786  row: o  column: 01
            High quality sequence stop: 734.
FEATURES              Location/Qualifiers
     source           1. .847
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4213848"
                      /tissue_type="glioblastoma with EGFR amplification"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /clone_lib="NCI_CGAP_Brn64"
                      /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                      Average insert size 1.57 kb. Constructed by Life
                      Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             1.5%;  Score 47;  DB 2;  Length 847;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-11;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGC 2934
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGC 82


RESULT 44
AA078268/c
LOCUS       AA078268                216 bp    mRNA    linear   EST 24-SEP-1999
DEFINITION  7H48E05 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone
            7H48E05, mRNA sequence.
ACCESSION   AA078268
VERSION     AA078268.1  GI:1837742
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,
            Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.
  TITLE     2006 expressed-sequence tags derived from human chromosome
            7-enriched cDNA libraries
  JOURNAL   Genome Res. 7 (3), 281-292 (1997)
   PUBMED   9074931
COMMENT     Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
            Tel: 3014020201
            Fax: 3014024735
            Email: egreen@nhgri.nih.gov
            Plate: 48  row: E  column: 05
            Seq primer: -21M13 (ABI).
FEATURES              Location/Qualifiers
     source           1. .216
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="7H48E05"
                      /sex="female"
                      /cell_line="HeLa cell line; ATCC"
                      /lab_host="E. coli strain DH5 alpha"
                      /clone_lib="Chromosome 7 HeLa cDNA Library"
                      /note="Vector: pAMP10; cDNA was generated from cytoplasmic
                      RNA using a mixture of random DNA hexamers and oligo(dT).
                      From this pool of cDNA, human chromosome 7-enriched cDNA
                      was isolated by direct cDNA selection using chromosome 7
                      genomic DNA (cosmids).  The resulting direct-selected cDNA
                      was cloned into a plasmid vector using a non-directional
                      uracil DNA glycosylase (UDG)-mediated cloning strategy."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGG 2933
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGG 37


RESULT 45
AI309360
LOCUS       AI309360                255 bp    mRNA    linear   EST 08-APR-1999
DEFINITION  tb51c03.x1 NCI_CGAP_Br15 Homo sapiens cDNA clone IMAGE:2057860 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI309360
VERSION     AI309360.1  GI:4004231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 255)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.
            student, Michael R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 317   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .255
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2057860"
                      /sex="female"
                      /tissue_type="adenocarcinoma"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Br15"
                      /note="Organ: breast; Vector: pAMP1; mRNA made from breast
                      adenocarcinoma tissue, cDNA made by oligo-dT  priming.
                      Directionally cloned.  Size-selected on agarose  gel,
                      average insert size 400 bp.  Primary library,
                      non-amplified."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        166 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 211


RESULT 46
BX475355
LOCUS       BX475355                313 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686M14176_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686M14176 5', mRNA sequence.
ACCESSION   BX475355
VERSION     BX475355.1  GI:31672638
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686M14176) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686M14176"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 5;  Length 313;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 92


RESULT 47
AW841960
LOCUS       AW841960                314 bp    mRNA    linear   EST 18-MAY-2000
DEFINITION  IL5-CN0024-100300-030-H12 CN0024 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW841960
VERSION     AW841960.1  GI:7935943
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 314)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL5-CN0024-100
            300-030-H12&t3=2000-03-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 297.
FEATURES             Location/Qualifiers
     source          1. .314
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CN0024"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 218


RESULT 48
CD514740/c
LOCUS       CD514740                319 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14394439 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30408431 5', mRNA sequence.
ACCESSION   CD514740
VERSION     CD514740.1  GI:31446458
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 319)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
            Plate: NDAM491  row: f  column: 24
            High quality sequence stop: 319.
FEATURES              Location/Qualifiers
     source           1. .319
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30408431"
                      /tissue_type="White Matter"
                      /dev_stage="Unknown"
                      /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                      /clone_lib="NIH_MGC_181"
                      /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                      (destroyed); Library is oligo-dT primed and directionally
                      cloned (EcoRV site is destroyed upon cloning). Average
                      insert size 1.42 kb. Library was constructed by
                      (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 6;  Length 319;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        116 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 71


RESULT 49
AA661680/c
LOCUS       AA661680                329 bp    mRNA    linear   EST 12-NOV-1997
DEFINITION  nu87b06.s1 NCI_CGAP_Alv1 Homo sapiens cDNA clone IMAGE:1217651
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA661680
VERSION     AA661680.1  GI:2615771
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
            Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 309.
FEATURES              Location/Qualifiers
     source           1. .329
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1217651"
                      /tissue_type="alveolar rhabdomyosarcoma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Alv1"
                      /note="Vector: pAMP10; mRNA made from alveolar
                      rhabdomyosarcoma, cDNA made by oligo-dT priming.
                      Non-directionally cloned. Size-selected on agarose gel,
                      average insert  size 600 bp. Reference:  Krizman et al.
                      (1996) Cancer Research 56:5380-5383."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 329;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 196


RESULT 50
BE061760/c
LOCUS       BE061760                341 bp    mRNA    linear   EST 09-JUN-2000
DEFINITION  RC1-BT0254-111099-012-a10 BT0254 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE061760
VERSION     BE061760.1  GI:8406410
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 341)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC1-BT0254-111
            099-012-a10&t3=1999-10-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 83
            High quality sequence stop: 139.
FEATURES              Location/Qualifiers
     source           1. .341
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="BT0254"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 2;  Length 341;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        166 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 121
```

```
RESULT 51
BF767274
LOCUS       BF767274                372 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  RC2-CN0084-091200-011-c12 CN0084 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF767274
VERSION     BF767274.1  GI:12115174
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC2&t2=RC2-CN0084-
            091200-011-c12&t3=2000-12-09&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 372.
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CN0084"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 2;  Length 372;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 226


RESULT 52
AQ201612/c
LOCUS       AQ201612                387 bp    DNA     linear   GSS 20-APR-1999
DEFINITION  RPCI11-59M4.TK RPCI-11 Homo sapiens genomic clone RPCI-11-59M4,
            genomic survey sequence.
ACCESSION   AQ201612
VERSION     AQ201612.1  GI:3609083
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 387)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-59M4.TJ
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7522563"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-59M4"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 387;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 244


RESULT 53
AQ134220/c
LOCUS       AQ134220                389 bp    DNA     linear   GSS 23-SEP-1998
DEFINITION  HS_3047_B1_A12_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3047 Col=23 Row=B, genomic survey
            sequence.
ACCESSION   AQ134220
VERSION     AQ134220.1  GI:3525586
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
```

```
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3047  row: B  column: 23
            Class: BAC ends
            High quality sequence stop: 389.
FEATURES              Location/Qualifiers
     source           1. .389
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=3047 Col=23 Row=B"
                      /sex="male"
                      /clone_lib="CIT Approved Human Genomic Sperm Library D"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        266 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 221


RESULT 54
AA503019/c
LOCUS       AA503019                405 bp    mRNA    linear   EST 15-AUG-1997
DEFINITION  ng20d04.s1 NCI_CGAP_Ov2 Homo sapiens cDNA clone IMAGE:929959
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA503019
VERSION     AA503019.1  GI:2237986
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1127   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 392.
FEATURES              Location/Qualifiers
     source           1. .405
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:929959"
                      /sex="female"
                      /tissue_type="ovary"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Ov2"
                      /note="Vector: pAMP10; mRNA made from invasive ovarian
                      tumor, cDNA made by oligo-dT priming.  Non-directionally
                      cloned. Size-selected on agarose gel, average insert  size
                      600 bp. Reference:  Krizman et al. (1996) Cancer Research
                      56:5380-5383."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        106 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 61


RESULT 55
AQ284177
LOCUS       AQ284177                406 bp    DNA     linear   GSS 27-APR-1999
DEFINITION  RPCI11-80L20.TV RPCI-11 Homo sapiens genomic clone RPCI-11-80L20,
            genomic survey sequence.
ACCESSION   AQ284177
VERSION     AQ284177.1  GI:3910495
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .406
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7530619"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-80L20"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /clone_lib="RPCI-11"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 406;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||
Db         25 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 70


RESULT 56
AW815323/c
LOCUS       AW815323                445 bp    mRNA    linear   EST 17-MAY-2000
DEFINITION  QV0-ST0215-060100-083-d08 ST0215 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW815323
VERSION     AW815323.1  GI:7908317
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV0-ST0215-060
            100-083-d08&t3=2000-01-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 445.
FEATURES             Location/Qualifiers
     source          1. .445
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ST0215"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 445;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        247 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 202


RESULT 57
AA504863
LOCUS       AA504863                 462 bp    mRNA    linear   EST 12-AUG-1997
DEFINITION  ab03b02.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
            IMAGE:839691 3' similar to contains Alu repetitive element;contains
            L1.t3 L1 repetitive element ;, mRNA sequence.
ACCESSION   AA504863
VERSION     AA504863.1  GI:2241023
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 399.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:839691"
                     /sex="mixed"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene fetal retina 937202"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Pooled
                     retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db         14 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 59


RESULT 58
AA128941/c
LOCUS       AA128941                 464 bp    mRNA    linear   EST 23-DEC-1997
DEFINITION  zo11c09.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens
            cDNA clone IMAGE:567376 3' similar to contains Alu repetitive
            element;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION   AA128941
VERSION     AA128941.1  GI:1688804
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            WARNING: There is evidence that suggests that the 384-well parent
            plate of this clone contains both human and mouse derived clones.
```

```
            Thus, the origin of this clone is uncertain. This caution should be
            kept in mind should you use this clone.

            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 706   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 317.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:4591775"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:567376"
                     /dev_stage="Ntera-2/RA+MI neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene neuroepithelium NT2RAMI 937234"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. NT2
                     (Ntera-2/c1.D1)  precursor cells induced with Retinoic
                     Acid for 1 week, followed by 3 weeks in mitotic inhibitors
                     (Replate #2). Average insert size: 1.1 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 1;  Length 464;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 161


RESULT 59
AQ383997
LOCUS       AQ383997                482 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-152K20.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-152K20,
            genomic survey sequence.
ACCESSION   AQ383997
VERSION     AQ383997.1  GI:4355020
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 482)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-152K20.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .482
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7558243"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-152K20"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 465


RESULT 60
AQ132449
LOCUS       AQ132449                507 bp    DNA     linear   GSS 23-SEP-1998
DEFINITION  HS_3003_A2_A08_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3003 Col=16 Row=A, genomic survey
            sequence.
ACCESSION   AQ132449
VERSION     AQ132449.1  GI:3509615
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3003  row: A  column: 16
            Class: BAC ends
            High quality sequence stop: 507.
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3003 Col=16 Row=A"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 507;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 170
```

```
RESULT 61
BX641230/c
LOCUS       BX641230                527 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686D19134_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686D19134 5', mRNA sequence.
ACCESSION   BX641230
VERSION     BX641230.1  GI:34475549
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686D19134) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686D19134"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 5;  Length 527;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 457


RESULT 62
BU960416
LOCUS       BU960416                550 bp    mRNA    linear   EST 21-OCT-2002
DEFINITION  AGENCOURT_10628034 NIH_MGC_141 Homo sapiens cDNA clone
            IMAGE:6739149 5', mRNA sequence.
ACCESSION   BU960416
VERSION     BU960416.1  GI:24189988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3071  row: o  column: 20
            High quality sequence stop: 436.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6739149"
                     /tissue_type="mixed (pool of 40 RNAs)"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_141"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
                     prepared from a pool of 40 cell line polyA+ RNAs (bladder
                     - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                     4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                     kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
                     ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
                     gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
                     used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.2-0.5
                     kb size fraction (other fractions present in NIH_MGC_142).
                     Library created in the laboratory of M. Brownstein (NIMH,
                     NIH). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 5;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 342


RESULT 63
CR960770/c
LOCUS       CR960770                564 bp    DNA     linear   GSS 06-JUN-2005
DEFINITION  Homo sapiens BAC end sequence of RZPDB737F022022D from genomic
            library (orig. Pieter J. de Jong library RPCI-11), genomic survey
            sequence.
ACCESSION   CR960770
VERSION     CR960770.1  GI:66991336
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Schwarz,F., Neubert,P., Schneider,D., Peters,M. and Korn,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2005) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH, Im Neuenheimer Feld 515, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDB737F022022D; RP11-59M4;
            derived from Pieter J. de Jong library RPCI-11;
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDB737F022022D
            RZPDLIB; (Human Genomic Set - RZPD 1.0) RZPD LIB No.737
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=737
```

```
            http://www.rzpd.de/products/genomicset/
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            Contact RZPD (customer.service@rzpd.de) for further information.
            Clone distribution: http://www.rzpd.de/products/genomicset/
            Seq-primer: T7 (TAA-TAC-GAC-TCA-CTA-TAG-GG)
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .564
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDB737F022022D"
                      /sex="Male"
                      /clone_lib="RPCI-11"
                      /note="Vector: pBACe3.6; RPCI-11 Human Male BAC Library"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 11;  Length 564;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        221 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 176


RESULT 64
CN265832
LOCUS       CN265832                 574 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424824545 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN265832
VERSION     CN265832.1  GI:47282246
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 574   Std Error: 0.00.
FEATURES              Location/Qualifiers
     source           1. .574
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /tissue_type="embryonic stem cells, embryoid bodies
                      derived from H1, H7 and H9 cells"
                      /clone_lib="GRN_EB"
                      /note="oligo dT primed, full-length enriched cDNA library
                      from embryoid body outgrowths derived from hES cell lines
                      H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                      conditions."
ORIGIN

  Query Match             1.5%;  Score 46;  DB 7;  Length 574;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        169 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 214


RESULT 65
B95171/c
LOCUS       B95171                   587 bp    DNA     linear   GSS 26-JUN-1998
DEFINITION  CIT-HSP-2172H8.TF CIT-HSP Homo sapiens genomic clone 2172H8,
            genomic survey sequence.
ACCESSION   B95171
VERSION     B95171.1  GI:2977508
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 587)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .587
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7103850"
                      /db_xref="taxon:9606"
                      /clone="2172H8"
                      /sex="Male"
                      /cell_type="Sperm"
                      /clone_lib="CIT-HSP"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        564 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 519


RESULT 66
AQ017813/c
LOCUS       AQ017813                 638 bp    DNA     linear   GSS 09-JUN-1998
DEFINITION  CIT-HSP-2307L15.TR CIT-HSP Homo sapiens genomic clone 2307L15,
            genomic survey sequence.
ACCESSION   AQ017813
VERSION     AQ017813.1  GI:3196549
KEYWORDS    GSS.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2307L15.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2307L15"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 9;  Length 638;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        551 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 506


RESULT 67
BX956845
LOCUS       BX956845                 648 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  DKFZp781H1492_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781H1492 5', mRNA sequence.
ACCESSION   BX956845
VERSION     BX956845.1  GI:43442084
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781H1492) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781H1492"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 5;  Length 648;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCC 2935
              |||||||||||||||||||||||||||||||||||||||||||||
Db         47 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCC 92


RESULT 68
BX641474/c
LOCUS       BX641474                 657 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686I14135_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686I14135 5', mRNA sequence.
ACCESSION   BX641474
VERSION     BX641474.1  GI:34475793
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686I14135) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686I14135"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
```

```
ORIGIN

  Query Match             1.5%;  Score 46;  DB 5;  Length 657;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 457


RESULT 69
CA430010
LOCUS       CA430010                687 bp    mRNA    linear   EST 07-NOV-2002
DEFINITION  UI-H-FH1-bfp-j-24-0-UI.s1 NCI_CGAP_FH1 Homo sapiens cDNA clone
            UI-H-FH1-bfp-j-24-0-UI 3', mRNA sequence.
ACCESSION   CA430010
VERSION     CA430010.1  GI:24792736
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 1-36, >AT_rich#Low_complexity (matched compliment)
            285-317, >AT_rich#Low_complexity (matched compliment) 369-658, >ALU
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FH1-bfp-j-24-0-UI"
                     /tissue_type="Cell Line"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FH1"
                     /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                     (Pharmacia) with a modified polylinker; Site_1: EcoR I;
                     Site_2: Not I; NCI_CGAP_FH1 is a normalized cDNA library
                     obtained from a cell line derived from grade I
                     chondrosarcoma tissue. The library was constructed and
                     normalized according to Bonaldo, Lennon and Soares, Genome
                     Research, 6:791-806, 1996. First strand cDNA synthesis was
                     primed with an oligo-dT primer containing a Not I site.
                     Double stranded cDNA was ligated to an EcoR I adaptor,
                     digested with Not I, and cloned directionally into
                     pT7T3-Pac vector. The oligonucleotide used to prime the
                     synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AGAATCCGGC. The cell line was provided by Dr. James Martin
                     from the University of Iowa.
                     TAG_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
                     Chondrosarcoma
                     TAG_LIB=UI-H-FH1
                     TAG_SEQ=AGAATCCGGC"
ORIGIN

  Query Match             1.5%;  Score 46;  DB 6;  Length 687;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 469


RESULT 70
BU564032
LOCUS       BU564032                258 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10371369 NIH_MGC_141 Homo sapiens cDNA clone
            IMAGE:6601889 5', mRNA sequence.
ACCESSION   BU564032
VERSION     BU564032.1  GI:22914332
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 258)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2840  row: d  column: 17
            High quality sequence stop: 171.
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6601889"
                     /tissue_type="mixed (pool of 40 RNAs)"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_141"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
                     prepared from a pool of 40 cell line polyA+ RNAs (bladder
                     - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                     4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                     kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
                     ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
                     gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
                     used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.2-0.5
                     kb size fraction (other fractions present in NIH_MGC_142).
                     Library created in the laboratory of M. Brownstein (NIMH,
                     NIH). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db        174 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 218
```

```
RESULT 71
BX953709
LOCUS       BX953709                287 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  DKFZp781C034_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781C034 5', mRNA sequence.
ACCESSION   BX953709
VERSION     BX953709.1  GI:43433046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 287)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp781C034) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .287
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781C034"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 287;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 285


RESULT 72
AQ485964/c
LOCUS       AQ485964                309 bp    DNA     linear   GSS 24-APR-1999
DEFINITION  RPCI-11-235I9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-235I9,
            genomic survey sequence.
ACCESSION   AQ485964
VERSION     AQ485964.1  GI:4668012
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7590056"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-235I9"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        209 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 165


RESULT 73
AQ381523
LOCUS       AQ381523                352 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-164A21.TV RPCI-11 Homo sapiens genomic clone RPCI-11-164A21,
            genomic survey sequence.
ACCESSION   AQ381523
VERSION     AQ381523.1  GI:4352546
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 352)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-164A21.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
```

```
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .352
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7562612"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-164A21"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 352;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db        145 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 189


RESULT 74
AQ485959/c
LOCUS       AQ485959                 408 bp    DNA     linear   GSS 24-APR-1999
DEFINITION  RPCI-11-235I5.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-235I5,
            genomic survey sequence.
ACCESSION   AQ485959
VERSION     AQ485959.1  GI:4668007
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-235I5.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7590052"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-235I5"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 408;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        306 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 262


RESULT 75
BU951869/c
LOCUS       BU951869                 415 bp    mRNA    linear   EST 21-OCT-2002
DEFINITION  in74a04.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6127759 3',
            mRNA sequence.
ACCESSION   BU951869
VERSION     BU951869.1  GI:24203621
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 415)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Sceare,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 337.
FEATURES             Location/Qualifiers
     source          1. .415
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6127759"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db         51 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 7
```

```
RESULT 76
AQ065412/c
LOCUS       AQ065412                417 bp    DNA     linear   GSS 04-AUG-1998
DEFINITION  HS_2224_B1_F05_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2224 Col=9 Row=L, genomic survey
            sequence.
ACCESSION   AQ065412
VERSION     AQ065412.1  GI:3380930
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2224  row: L  column: 9
            Class: BAC ends
            High quality sequence stop: 417.
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2224 Col=9 Row=L"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        414 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 370


RESULT 77
AA085683/c
LOCUS       AA085683                424 bp    mRNA    linear   EST 01-DEC-1996
DEFINITION  zn53d01.s1 Stratagene muscle 937209 Homo sapiens cDNA clone
            IMAGE:561889 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA085683
VERSION     AA085683.1  GI:1629145
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2224   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 393.
FEATURES             Location/Qualifiers
     source          1. .424
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:4595179"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:561889"
                     /tissue_type="muscle"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene muscle 937209"
                     /note="Organ: skeletal muscle; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
                     Primer: Oligo dT. Skeletal muscle from patient with
                     malignant hyperthermia.  Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 1;  Length 424;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
              |||||||||||||||||||||||||||||||||||||||||||||
Db        222 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 178


RESULT 78
CB132267/c
LOCUS       CB132267                436 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0182497 L11SNU354s1 Homo sapiens cDNA clone L11SNU354s1-1-E08
            5', mRNA sequence.
ACCESSION   CB132267
VERSION     CB132267.1  GI:28097454
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 1  row: E  column: 08
```

```
            High quality sequence stop: 436.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L11SNU354s1-1-E08"
                     /sex="M"
                     /tissue_type="Liver"
                     /cell_type="Polygonal"
                     /cell_line="SNU-354"
                     /lab_host="Top10F'"
                     /clone_lib="L11SNU354s1"
                     /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library. After analyzing and
                     sequencing about 2,000 - 3,000 colonies in original cDNA
                     library, the abundant cDNAs were selected and amplified by
                     PCR reaction using vector region primer including T7
                     promotor as 5' primer and N(dT)14 as 3' primer. The PCR
                     products were used as template for synthesis of
                     biotinylated single stranded RNA by in vitro transcription
                     reaction. The synthesized RNA probes were hybridized with
                     antisense single stranded cDNAs prepared from original
                     liberary and incubated with avidin-gel. After removing
                     DNA-RNA hybrids by centrifuge, the subtracted cDNA
                     libraries were constructed by transformaion of the
                     remaining DNA into competent cells E. coli Top10F' with
                     electroporation method."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 6;  Length 436;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db        233 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 189


RESULT 79
AQ009550
LOCUS       AQ009550                439 bp    DNA     linear   GSS 27-JUN-1998
DEFINITION  CIT-HSP-2285K1.TRB CIT-HSP Homo sapiens genomic clone 2285K1,
            genomic survey sequence.
ACCESSION   AQ009550
VERSION     AQ009550.1  GI:3128935
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7148812"
                     /db_xref="taxon:9606"
                     /clone="2285K1"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 439;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
              |||||||||||||||||||||||||||||||||||||||||||||
Db        282 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 326


RESULT 80
AQ554616/c
LOCUS       AQ554616                447 bp    DNA     linear   GSS 28-MAY-1999
DEFINITION  RPCI-11-409A7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-409A7,
            genomic survey sequence.
ACCESSION   AQ554616
VERSION     AQ554616.1  GI:4913793
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-409A7.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Homo sapiens"
```

```
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7656678"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-409A7"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 447;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        182 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 138


RESULT 81
AQ170811
LOCUS       AQ170811                 452 bp    DNA     linear   GSS 16-OCT-1998
DEFINITION  HS_3071_A1_H09_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3071 Col=17 Row=O, genomic survey
            sequence.
ACCESSION   AQ170811
VERSION     AQ170811.1  GI:3568178
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3071  row: O  column: 17
            Class: BAC ends
            High quality sequence stop: 452.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3071 Col=17 Row=O"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 452;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        201 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 245


RESULT 82
BG944968
LOCUS       BG944968                 469 bp    mRNA    linear   EST 15-JAN-2003
DEFINITION  ax57a12.x1 Hembase; Erythroid Progenitor Cells (LCB:ax library)
            Homo sapiens cDNA clone ax57a12 random, mRNA sequence.
ACCESSION   BG944968
VERSION     BG944968.1  GI:14344340
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene expression in proliferating human erythroid cells
  JOURNAL   Genomics 59 (2), 168-177 (1999)
   PUBMED   10409428
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 57  row: a  column: 12
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .469
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ax57a12"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71++++"
                     /lab_host="SOLR"
                     /clone_lib="Hembase; Erythroid Progenitor Cells (LCB:ax
                     library)"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
                     Site_2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        132 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 176


RESULT 83
```

```
H63465/c
LOCUS       H63465                  472 bp    mRNA    linear   EST 11-OCT-1995
DEFINITION  yr53e06.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:209026 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   H63465
VERSION     H63465.1  GI:1018266
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1442
            High quality sequence stops: 426
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1442   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3781867"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:209026"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 8;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 238


RESULT 84
BX954324
LOCUS       BX954324                494 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  DKFZp781B2254_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781B2254 5', mRNA sequence.
ACCESSION   BX954324
VERSION     BX954324.1  GI:43434818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Heinrich-
            Heine-University, Duesseldorf/Germany) within the cDNA sequencing
            consortium of the German Genome Project. No s1 sequence available.
            This clone (DKFZp781B2254) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp781B2254"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="781 (synonym: hlcc4)"
                     /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 494;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         56 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 100


RESULT 85
AQ038737/c
LOCUS       AQ038737                503 bp    DNA     linear   GSS 11-JUL-1998
DEFINITION  CIT-HSP-2325O9.TV CIT-HSP Homo sapiens genomic clone 2325O9,
            genomic survey sequence.
ACCESSION   AQ038737
VERSION     AQ038737.1  GI:3304569
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2325O9.TRB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
```

```
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .503
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="232509"
                      /sex="Male"
                      /cell_type="Sperm"
                      /clone_lib="CIT-HSP"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 503;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        268 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 224


RESULT 86
CB142858/c
LOCUS       CB142858                529 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0196755 L11SNU354s1 Homo sapiens cDNA clone L11SNU354s1-22-H11
            5', mRNA sequence.
ACCESSION   CB142858
VERSION     CB142858.1  GI:28119424
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 22  row: H  column: 11
            High quality sequence stop: 529.
FEATURES              Location/Qualifiers
     source           1. .529
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="L11SNU354s1-22-H11"
                      /sex="M"
                      /tissue_type="Liver"
                      /cell_type="Polygonal"
                      /cell_line="SNU-354"
                      /lab_host="Top10F'"
                      /clone_lib="L11SNU354s1"
                      /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                      Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                      bacterial alkaline phosphatase (BAP) and then decapped
                      with tabacco acid pyrophosphatase (TAP). The decapped
                      intact mRNA was ligated with DNA-RNA linker including
                      EcoRI site by treatment of T4 RNA ligase and the first
                      strand cDNA was synthesized from oligo dT-selected mRNA by
                      priming with dT-tailed vector. The dT-tailed vector was
                      adjusted to have about 60nt. The cDNA vector was
                      circularized with E. coli DNA ligase after digestion of
                      EcoRI which site is also included in vector. An RNA strand
                      converted to a DNA strand by Okayama-Berg method. The
                      obtained cDNA vectors were used for transformation of
                      competent cells E. coli Top10F' by electroporation method.
                      The cDNA libraries constructed by this method are
                      full-length enriched cDNA library. After analyzing and
                      sequencing about 2,000 - 3,000 colonies in original cDNA
                      library, the abundant cDNAs were selected and amplified by
                      PCR reaction using vector region primer including T7
                      promotor as 5' primer and N(dT)14 as 3' primer. The PCR
                      products were used as template for synthesis of
                      biotinylated single stranded RNA by in vitro transcription
                      reaction. The synthesized RNA probes were hybridized with
                      antisense single stranded cDNAs prepared from original
                      liberary and incubated with avidin-gel. After removing
                      DNA-RNA hybrids by centrifuge, the subtracted cDNA
                      libraries were constructed by transformaion of the
                      remaining DNA into competent cells E. coli Top10F' with
                      electroporation method."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 6;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db        233 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 189


RESULT 87
AQ333702/c
LOCUS       AQ333702                529 bp    DNA     linear   GSS 06-MAR-1999
DEFINITION  HS_5013_A2_F10_T7 RPCI11 Human Male BAC Library Homo sapiens
            genomic clone Plate=589 Col=20 Row=K, genomic survey sequence.
ACCESSION   AQ333702
VERSION     AQ333702.1  GI:4131189
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 589  row: K  column: 20
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 529.
FEATURES              Location/Qualifiers
     source           1. .529
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
```

```
                     /clone="Plate=589 Col=20 Row=K"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
              |||||||||||||||||||||||||||||||||||||||||||||
Db        214 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 170


RESULT 88
AU158612/c
LOCUS       AU158612                 568 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU158612 PLACE3 Homo sapiens cDNA clone PLACE3000425 3', mRNA
            sequence.
ACCESSION   AU158612
VERSION     AU158612.1  GI:11020133
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
            Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
            Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
            Masuho,Y., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE3000425"
                     /tissue_type="placenta"
                     /clone_lib="PLACE3"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 1;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         47 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3


RESULT 89
BP872324/c
LOCUS       BP872324                 599 bp    mRNA    linear   EST 27-JUL-2004
DEFINITION  BP872324 Sugano cDNA library, embryonal kidney Homo sapiens cDNA
            clone HKR06243, mRNA sequence.
ACCESSION   BP872324
VERSION     BP872324.1  GI:50717491
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Nakai,K. and Sugano,S.
  TITLE     Database of Transcriptional Start Sites (DBTSS) for Comparative
            Studies of the Promoters of Human and Mouse Genes
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HKR06243"
                     /tissue_type="embryonal kidney"
                     /cell_line="293"
                     /clone_lib="Sugano cDNA library, embryonal kidney"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 3;  Length 599;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db        233 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 189


RESULT 90
AG071092
LOCUS       AG071092                 627 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-062C22.R, genomic survey sequence.
ACCESSION   AG071092
VERSION     AG071092.1  GI:16622894
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 627)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
```

```
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-062C22.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 627;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        225 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 269
.

RESULT 91
AG116078/c
LOCUS       AG116078                640 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-123I03.R, genomic survey sequence.
ACCESSION   AG116078
VERSION     AG116078.1  GI:16736597
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 640)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-123I03.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 640;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||||
Db        386 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 342

RESULT 92
AQ540193/c
LOCUS       AQ540193                641 bp    DNA     linear   GSS 19-MAY-1999
DEFINITION  RPCI-11-345P21.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-345P21, genomic survey sequence.
ACCESSION   AQ540193
VERSION     AQ540193.1  GI:4870723
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-345P21.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7632476"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-345P21"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 641;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         492 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 448


RESULT 93
AG174909/c
LOCUS       AG174909                680 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-045K01.T7, genomic survey
            sequence.
ACCESSION   AG174909
VERSION     AG174909.1  GI:16704589
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
```

```
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 680)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-045K01.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 680;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        118 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 74


RESULT 94
CN274316
LOCUS       CN274316                712 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600023281 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN274316
VERSION     CN274316.1  GI:47290730
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 712)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 712   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .712
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 7;  Length 712;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        319 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 363


RESULT 95
AQ349458
LOCUS       AQ349458                722 bp    DNA     linear   GSS 07-MAY-1999
DEFINITION  RPCI11-118J16.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-118J16,
            genomic survey sequence.
ACCESSION   AQ349458
VERSION     AQ349458.1  GI:4174354
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 722)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-118J16.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7545159"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-118J16"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 722;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||||
Db        516 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 560


RESULT 96
AG009127
LOCUS       AG009127                726 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: T172XN, genomic survey
            sequence.
ACCESSION   AG009127 AG003036
VERSION     AG009127.1  GI:3289113
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 726)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
COMMENT     On Feb 5, 1999 this sequence version replaced gi:2665425.
            AG003036: Submitted (06-Dec-1997).
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="T172XN"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 726;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
              |||||||||||||||||||||||||||||||||||||||||||||
Db        179 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 223


RESULT 97
AG009138/c
LOCUS       AG009138                731 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: T172XN, genomic survey
            sequence.
ACCESSION   AG009138 AG003047
VERSION     AG009138.1  GI:3289124
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 731)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
COMMENT     On Feb 5, 1999 this sequence version replaced gi:2665436.
            AG003047: Submitted (06-Dec-1997).
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="T172XN"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 731;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
              |||||||||||||||||||||||||||||||||||||||||||||
Db        274 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 230


RESULT 98
AG009129
LOCUS       AG009129                741 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: T172XN, genomic survey
            sequence.
ACCESSION   AG009129 AG003038
VERSION     AG009129.1  GI:3289115
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1998)
REFERENCE   2  (bases 1 to 741)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1998) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
COMMENT     On Feb 5, 1999 this sequence version replaced gi:2665427.
            AG003038: Submitted (06-Dec-1997).
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="T172XN"
ORIGIN

  Query Match             1.4%;  Score 45;  DB 10;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
              |||||||||||||||||||||||||||||||||||||||||||||
Db        381 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 425


RESULT 99
BX457023
LOCUS       BX457023                754 bp    mRNA    linear   EST 06-MAY-2004
```

```
DEFINITION  BX457023 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP005YD12
            5-PRIME, mRNA sequence.
ACCESSION   BX457023
VERSION     BX457023.2  GI:47069656
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 22, 2003 this sequence version replaced gi:31032820.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 5067.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0CAP005DB06QP1&c=5067.r.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP005YD12"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 754;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        622 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 666


RESULT 100
BG260049/c
LOCUS       BG260049                768 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602371633F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4479632 5',
            mRNA sequence.
ACCESSION   BG260049
VERSION     BG260049.1  GI:12769865
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10312  row: a  column: 09
            High quality sequence stop: 700.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4479632"
                     /tissue_type="transitional cell papilloma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_93"
                     /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 2;  Length 768;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 175


RESULT 101
BX411269/c
LOCUS       BX411269                781 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX411269 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF021YC03 3-PRIME, mRNA sequence.
ACCESSION   BX411269
VERSION     BX411269.2  GI:46955683
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30767155.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 8085.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAI011ZB10_CS01034_2&c=8085.r
            .
FEATURES             Location/Qualifiers
     source          1. .781
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF021YC03"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
```

enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 781;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        723 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 679
```

RESULT 102
CD242479
LOCUS CD242479 830 bp mRNA linear EST 22-MAY-2003
DEFINITION AGENCOURT_14120572 NIH_MGC_179 Homo sapiens cDNA clone IMAGE:30384600 5', mRNA sequence.
ACCESSION CD242479
VERSION CD242479.1 GI:31002943
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM450 row: f column: 01
High quality sequence stop: 493.
FEATURES Location/Qualifiers
source 1. .830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30384600"
/tissue_type="Pituitary"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_179"
/note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.4%; Score 45; DB 6; Length 830;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        496 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 540
```

RESULT 103
BU959380/c
LOCUS BU959380 856 bp mRNA linear EST 21-OCT-2002
DEFINITION AGENCOURT_10622545 NIH_MGC_127 Homo sapiens cDNA clone IMAGE:6737609 5', mRNA sequence.
ACCESSION BU959380
VERSION BU959380.1 GI:24188952
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3067 row: o column: 16
High quality sequence stop: 431.
FEATURES Location/Qualifiers
source 1. .856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6737609"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_127"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidnney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 856;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
              |||||||||||||||||||||||||||||||||||||||||||||
Db        203 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 159
```

RESULT 104
BZ771376/c
LOCUS BZ771376 874 bp DNA linear GSS 13-MAR-2003
DEFINITION mcs76c11.g10 HFOSMID005 Homo sapiens genomic, genomic survey sequence.
ACCESSION BZ771376
VERSION BZ771376.1 GI:28945045
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 874)
  AUTHORS   Cook,L., Delehaunty,K., Fewell,G., Fulton,L., Magrini,V.,
            Mardis,E., Miner,T., Nash,W., Williams,D. and Wilson,R.K.
  TITLE     Homo sapiens Fosmid End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Plate: mcs76  row: c  column: 11
            Class: fosmid ends
            High quality sequence start: 11
            High quality sequence stop: 518.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="HFOSMID005"
                     /note="Vector: pcc01fos; Site_1: Eco72i; Human whole
                     genome fosmid library was prepared at Washington
                     University Genome Sequencing Center. DNA was sheared for
                     blunt-ended ligation into pcc01fos inducible vector. DNA
                     was ordered from Coriell Cell Repository's DNA
                     polymorphism discovery resource."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 9;  Length 874;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||
Db        137 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 93


RESULT 105
CF596843/c
LOCUS       CF596843                948 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  AGENCOURT_15668731 NICHD_Hs_Ov1 Homo sapiens cDNA clone
            IMAGE:30705007 5', mRNA sequence.
ACCESSION   CF596843
VERSION     CF596843.1  GI:36353710
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gregory F. Erickson, Ph.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM254  row: p  column: 08
            High quality sequence stop: 432.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30705007"
                     /tissue_type="Ovary"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NICHD_Hs_Ov1"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Library is oligo-dT primed
                     and directionally cloned. Granulosa lutein cells aspirated
                     from preovulatory folicles of normal cycling women
                     undergoing ovulation induction for infertility due to male
                     factor and normal doners. The cells were from follicles
                     stimulated with Lupron, FSH and hCG. 5' and 3' adaptors
                     were used in cloning as follows: 5' adaptor sequence:
                     5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 2.23
                     kb (range 1.0-4.5 kb). 14/15 colonies contained inserts by
                     PCR. This library was enriched for full-length clones and
                     was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 6;  Length 948;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db        440 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 396


RESULT 106
BQ896885/c
LOCUS       BQ896885                988 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8073985 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6087057
            5', mRNA sequence.
ACCESSION   BQ896885
VERSION     BQ896885.1  GI:22288899
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2323  row: a  column: 10
            High quality sequence stop: 595.
FEATURES             Location/Qualifiers
     source          1. .988
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6087057"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_102"
                     /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of  California, Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies). Note: this is a NIH_MGC Library."
```

```
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 988;
  Best Local Similarity   100.0%;  Pred. No. 6e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              |||||||||||||||||||||||||||||||||||||||||||||
Db         53 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 9


RESULT 107
BX404721
LOCUS       BX404721               1123 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX404721 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP005YD12
            5-PRIME, mRNA sequence.
ACCESSION   BX404721
VERSION     BX404721.2  GI:46925367
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1123)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30639134.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 5067.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AAW13ZE04QP1&c=5067.r.
FEATURES             Location/Qualifiers
     source          1. .1123
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP005YD12"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             1.4%;  Score 45;  DB 5;  Length 1123;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        622 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 666


RESULT 108
BG959135
LOCUS       BG959135               203 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  PM4-CT0806-180301-003-g01 CT0806 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG959135
VERSION     BG959135.1  GI:14377306
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 203)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM4&t2=PM4-CT0806-
            180301-003-g01&t3=2001-03-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 127.
FEATURES             Location/Qualifiers
     source          1. .203
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0806"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 203;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 104


RESULT 109
BG010132/c
LOCUS       BG010132               219 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  MR3-GN0185-041200-015-f11 GN0185 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG010132
VERSION     BG010132.1  GI:12457017
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 219)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
```

```
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-GN0185-
            041200-015-f11&t3=2000-12-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 10
            High quality sequence stop: 219.
FEATURES             Location/Qualifiers
     source          1. .219
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0185"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 219;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||
Db        218 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 175


RESULT 110
BF893386
LOCUS       BF893386                281 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  QV3-MT0129-111100-427-h11 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF893386
VERSION     BF893386.1  GI:12284845
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 281)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-MT0129-
            111100-427-h11&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 5
            High quality sequence stop: 281.
FEATURES             Location/Qualifiers
     source          1. .281
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0129"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              ||||||||||||||||||||||||||||||||||||||||||||
Db         67 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 110


RESULT 111
BH609712
LOCUS       BH609712                281 bp    DNA     linear   GSS 18-DEC-2001
DEFINITION  HIV18E03 SupT1 HIV-I in vivo integration lines Homo sapiens genomic
            clone HIV18E03, genomic survey sequence.
ACCESSION   BH609712
VERSION     BH609712.1  GI:17922321
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 281)
  AUTHORS   Schroder,A.R.W., Shinn,P., Chen,H., Berry, C., Ecker,J.R. and
            Bushman,F.
  TITLE     Favored Sites for HIV-1 Integration in the Human Genome
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Frederic Bushman
            Salk Institute Infectious Disease Laboratory
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1630
            Fax: 858 554 0341
            Email: bushman@salk.edu
            Class: PCR with specific primers.
FEATURES             Location/Qualifiers
     source          1. .281
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="HIV18E03"
                     /cell_line="SupT1"
                     /clone_lib="SupT1 HIV-I in vivo integration lines"
                     /note="A human T-cell line (SupT1) was infected with an
                     HIV-based vector. DNA was isolated and cleaved with
                     restriction enzymes; linkers were ligated onto the cleaved
                     DNA and DNAs were amplified using one primer that bound to
```

```
                the linker DNA and one that bound to the HIV cDNA.
                Junctions between integrated HIV proviruses and cellular
                DNA were cloned and sequenced."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              |||||||||||||||||||||||||||||||||||||||||||||
Db        148 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 191


RESULT 112
AA493894/c
LOCUS       AA493894                294 bp    mRNA    linear   EST 20-AUG-1997
DEFINITION  nh07b12.s1 NCI_CGAP_Thy1 Homo sapiens cDNA clone IMAGE:943583
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AA493894
VERSION     AA493894.1  GI:2223735
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 679   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 256.
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:943583"
                     /tissue_type="thyroid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Thy1"
                     /note="Vector: pAMP10; mRNA made from invasive thyroid
                     tumor, cDNA made by oligo-dT priming.  Non-directionally
                     cloned. Size-selected on agarose gel, average insert  size
                     600 bp. Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         74 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 31


RESULT 113
AI382825/c
LOCUS       AI382825                301 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  ta72h10.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:2049667 3' similar to contains Alu repetitive
            element;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI382825
VERSION     AI382825.1  GI:4195606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 301)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 923   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 287.
FEATURES             Location/Qualifiers
     source          1. .301
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2049667"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 301;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||
Db         94 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 51


RESULT 114
BF879334
LOCUS       BF879334                301 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  IL3-ET0111-231100-337-C04 ET0111 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF879334
VERSION     BF879334.1  GI:12269464
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 301)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
```

```
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-ET0111-
            231100-337-C04&t3=2000-11-23&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .301
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0111"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 301;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 226


RESULT 115
BF914587/c
LOCUS       BF914587                325 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  IL3-UT0114-011200-362-C02 UT0114 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF914587
VERSION     BF914587.1  GI:12306045
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-UT0114-
            011200-362-C02&t3=2000-12-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0114"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        224 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 181


RESULT 116
AQ059204/c
LOCUS       AQ059204                337 bp    DNA     linear   GSS 31-JUL-1998
DEFINITION  CIT-HSP-2349B18.TR CIT-HSP Homo sapiens genomic clone 2349B18,
            genomic survey sequence.
ACCESSION   AQ059204
VERSION     AQ059204.1  GI:3361541
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 337)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2349B18.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .337
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2349B18"
                     /sex="Male"
                     /cell_type="Sperm"
```

```
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        192 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 149


RESULT 117
N63149/c
LOCUS       N63149                  338 bp    mRNA    linear   EST 01-MAR-1996
DEFINITION  yz37e10.s1 Morton Fetal Cochlea Homo sapiens cDNA clone
            IMAGE:285258 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N63149
VERSION     N63149.1  GI:1210978
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 338)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 382.
FEATURES             Location/Qualifiers
     source          1. .338
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3892375"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:285258"
                     /tissue_type="cochlea"
                     /dev_stage="16-22 week fetus"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /clone_lib="Morton Fetal Cochlea"
                     /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
                     Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
                     unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
                     37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
                     XR Vector. Library constructed by N. Robertson, C. Morton.
                     ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 8;  Length 338;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         44 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 1


RESULT 118
AQ092587/c
LOCUS       AQ092587                352 bp    DNA     linear   GSS 27-AUG-1998
DEFINITION  HS_3003_B2_A09_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3003 Col=18 Row=B, genomic survey
            sequence.
ACCESSION   AQ092587
VERSION     AQ092587.1  GI:3464034
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 352)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3003  row: B  column: 18
            Class: BAC ends
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1. .352
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3003 Col=18 Row=B"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 352;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        165 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 122


RESULT 119
AI367551/c
LOCUS       AI367551                353 bp    mRNA    linear   EST 15-FEB-1999
DEFINITION  qv99e05.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:1989728 3'
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AI367551
VERSION     AI367551.1  GI:4137296
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 353)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1947   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1. .353
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1989728"
                     /tissue_type="moderately-differentiated endometrial
                     adenocarcinoma, 3 pooled tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ut2"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.85 kb. Life Technologies catalog #:
                     11539-012"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 353;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db         45 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 2


RESULT 120
AA654781/c
LOCUS       AA654781                365 bp    mRNA    linear   EST 04-NOV-1997
DEFINITION  nt73g04.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204182
            similar to contains Alu repetitive element;contains element PTR5
            MER4 repetitive element ;, mRNA sequence.
ACCESSION   AA654781
VERSION     AA654781.1  GI:2590935
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 321.
FEATURES             Location/Qualifiers
     source          1. .365
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1204182"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr3"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from
                     5,000-10,000 microdissected cells
                     histologically-determined to be fully malignant prostate
                     cancer cells.  Double-stranded cDNA was ligated to EcoRI
                     adaptors, 5 cycles of PCR applied to the cDNA with an
                     adaptor-specific primer, and the resulting PCR product
                     subcloned into pAMP10 by the UDG-cloning method (Life
                     Technologies).  Average insert size is 600 bp. NOTE: Not
                     directionally cloned.  This library was constructed by
                     David Krizman."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db         74 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 31


RESULT 121
AQ061311
LOCUS       AQ061311                377 bp    DNA     linear   GSS 31-JUL-1998
DEFINITION  CIT-HSP-2349M9.TF CIT-HSP Homo sapiens genomic clone 2349M9,
            genomic survey sequence.
ACCESSION   AQ061311
VERSION     AQ061311.1  GI:3363223
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 377)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2349M9.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .377
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2349M9"
```

```
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 377;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
              ||||||||||||||||||||||||||||||||||||||||||||
Db        209 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 252


RESULT 122
AA653916/c
LOCUS       AA653916                 378 bp    mRNA    linear   EST 04-NOV-1997
DEFINITION  nt80a05.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204784
            similar to contains Alu repetitive element;contains element MSR1
            repetitive element ;, mRNA sequence.
ACCESSION   AA653916
VERSION     AA653916.1  GI:2590070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .378
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1204784"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr3"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from
                     5,000-10,000 microdissected cells
                     histologically-determined to be fully malignant prostate
                     cancer cells.  Double-stranded cDNA was ligated to EcoRI
                     adaptors, 5 cycles of PCR applied to the cDNA with an
                     adaptor-specific primer, and the resulting PCR product
                     subcloned into pAMP10 by the UDG-cloning method (Life
                     Technologies).  Average insert size is 600 bp. NOTE: Not
                     directionally cloned.  This library was constructed by
                     David Krizman."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db         50 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7


RESULT 123
AQ239365/c
LOCUS       AQ239365                 381 bp    DNA     linear   GSS 30-SEP-1998
DEFINITION  CIT-HSP-2383D24.TR.1 CIT-HSP Homo sapiens genomic clone 2383D24,
            genomic survey sequence.
ACCESSION   AQ239365
VERSION     AQ239365.1  GI:3671656
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2383D24"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              |||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 168


RESULT 124
BU561388
LOCUS       BU561388                 382 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10278700 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6592405
            5', mRNA sequence.
ACCESSION   BU561388
VERSION     BU561388.1  GI:22911684
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 382)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2815  row: i  column: 13
            High quality sequence stop: 381.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6592405"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||
Db        233 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 276


RESULT 125
AA728990/c
LOCUS       AA728990               389 bp    mRNA    linear   EST 06-JAN-1998
DEFINITION  nw22g06.s1 NCI_CGAP_GCB0 Homo sapiens cDNA clone IMAGE:1241242 3'
            similar to contains Alu repetitive element;contains L1.t1 L1
            repetitive element ;, mRNA sequence.
ACCESSION   AA728990
VERSION     AA728990.1  GI:2750349
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 367.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1241242"
                     /tissue_type="germinal center B-cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="NCI_CGAP_GCB0"
                     /note="Organ: tonsil; Vector: Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Germinal center B-cells Library constructed by
                     Dr. L. Staudt (NCI). 5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size:  1.1 kb."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||
Db         58 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 15


RESULT 126
H44630
LOCUS       H44630                 409 bp    mRNA    linear   EST 31-JUL-1995
DEFINITION  yp19g10.s1 Soares breast 3NbHBst Homo sapiens cDNA clone
            IMAGE:187938 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   H44630
VERSION     H44630.1  GI:920682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 409)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1706
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1706   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 345.
FEATURES             Location/Qualifiers
     source          1. .409
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3818835"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:187938"
                     /sex="Female"
                     /dev_stage="adult"
```

```
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares breast 3NbHBst"
                     /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                     Library went through one round of normalization to a Cot =
                     20. Library constructed by Bento Soares and M.Fatima
                     Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 8;  Length 409;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        314 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 357
-

RESULT 127
AQ605062/c
LOCUS       AQ605062                412 bp    DNA     linear   GSS 10-JUN-1999
DEFINITION  HS_2119_B1_G05_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2119 Col=9 Row=N, genomic survey
            sequence.
ACCESSION   AQ605062
VERSION     AQ605062.1  GI:5065056
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2119  row: N  column: 9
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 412.
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2119 Col=9 Row=N"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db         88 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 45


RESULT 128
CR769029/c
LOCUS       CR769029                421 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp468I1312_r1 468 (synonym: phrt1) Pongo pygmaeus cDNA clone
            DKFZp468I1312 5', mRNA sequence.
ACCESSION   CR769029
VERSION     CR769029.1  GI:52612739
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
            ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp468I1312
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp468I1312"
                     /tissue_type="heart"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="468 (synonym: phrt1)"
                     /note="Vector: pSport1_SfiI; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        145 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 102


RESULT 129
N20066
LOCUS       N20066                  426 bp    mRNA    linear   EST 15-DEC-1995
DEFINITION  yx28g04.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:263094 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N20066
VERSION     N20066.1  GI:1124733
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
```

```
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 404
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2002   Std Error: 0.00
            Seq primer: m13 -40 forward
            High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3872736"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:263094"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 8;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        380 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 423


RESULT 130
CR546695/c
LOCUS       CR546695                435 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  DKFZp470L1324_r1 470 (synonym: pliv1) Pongo pygmaeus cDNA clone
            DKFZp470L1324 5', mRNA sequence.
ACCESSION   CR546695
VERSION     CR546695.1  GI:49899216
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 435)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Wambutt,R., Heubner,D., Mewes,H.W., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa
            (Berlin/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp470L1324) is available at
            the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
            Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
            clone@rzpd.de Further information about the clone and the
            sequencing project is available at
            http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .435
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp470L1324"
                     /tissue_type="liver"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="470 (synonym: pliv1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 435;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        178 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 135


RESULT 131
B60059/c
LOCUS       B60059                  438 bp    DNA     linear   GSS 21-JUN-1998
DEFINITION  CIT-HSP-385O15.TFB CIT-HSP Homo sapiens genomic clone 385O15,
            genomic survey sequence.
ACCESSION   B60059
VERSION     B60059.1  GI:2614777
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-385O15.TRB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:5378896"
                     /db_xref="taxon:9606"
                     /clone="385O15"
                     /sex="Male"
                     /cell_type="Sperm"
```

```
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 438;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 393


RESULT 132
AL713064
LOCUS       AL713064                447 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686C1395_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686C1395 5', mRNA sequence.
ACCESSION   AL713064
VERSION     AL713064.1  GI:19696420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,W., Weil,B. and
            Wiemann,S.
  TITLE     EST (Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and
            Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686C1395) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686C1395"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 447;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        264 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 307


RESULT 133
AI469968
LOCUS       AI469968                448 bp    mRNA    linear   EST 14-APR-1999
DEFINITION  tj89c03.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2148676 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AI469968
VERSION     AI469968.1  GI:4332058
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 710   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2148676"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        381 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 424


RESULT 134
AQ705037/c
LOCUS       AQ705037                448 bp    DNA     linear   GSS 07-JUL-1999
DEFINITION  HS_5521_B2_H07_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1097 Col=14 Row=P, genomic survey sequence.
ACCESSION   AQ705037
VERSION     AQ705037.1  GI:5414463
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
```

```
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1097  row: P  column: 14
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=1097 Col=14 Row=P"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 348


RESULT 135
AU157238/c
LOCUS       AU157238                466 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU157238 PLACE1 Homo sapiens cDNA clone PLACE1006883 3', mRNA
            sequence.
ACCESSION   AU157238
VERSION     AU157238.1  GI:11018759
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T.,
            Sugano,S. and Isogai,T.).
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
            Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M.,
            Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1006883"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         44 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 1


RESULT 136
BX111733
LOCUS       BX111733                466 bp    mRNA    linear   EST 07-FEB-2003
DEFINITION  BX111733 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGp998D133381 ;
            IMAGE:1340508, mRNA sequence.
ACCESSION   BX111733
VERSION     BX111733.1  GI:27837211
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998D133381.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998D133381 ; IMAGE:1340508"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
```

primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db         83 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 126
```

RESULT 137
BE674703/c
LOCUS BE674703 468 bp mRNA linear EST 08-SEP-2000
DEFINITION 7e94e04.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:3292830 3' similar to contains Alu repetitive element;contains element MIR repetitive element ;, mRNA sequence.
ACCESSION BE674703
VERSION BE674703.1 GI:10035325
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 468)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 454.
FEATURES Location/Qualifiers
source 1. .468
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3292830"
/tissue_type="B-cell, chronic lymphotic leukemia"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATTGCTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 1.4%; Score 44; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              ||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 198
```

RESULT 138
T84449/c
LOCUS T84449 472 bp mRNA linear EST 16-MAR-1995
DEFINITION yd32d04.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:109927 5' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION T84449
VERSION T84449.1 GI:712737
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 472)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 862
High quality sequence stops: 363 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 862 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 363.
FEATURES Location/Qualifiers
source 1. .472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:465544"
/db_xref="taxon:9606"
/clone="IMAGE:109927"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||
Db        138 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 95
```

RESULT 139
BM983814/c

```
LOCUS       BM983814                473 bp    mRNA    linear   EST 20-FEB-2003
DEFINITION  UI-CF-DU1-aay-c-24-0-UI.s1 UI-CF-DU1 Homo sapiens cDNA clone
            UI-CF-DU1-aay-c-24-0-UI 3', mRNA sequence.
ACCESSION   BM983814
VERSION     BM983814.1  GI:19608703
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  PUBMED    8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
             The following repetitive elements were found in this cDNA
            sequence: 11-219, >ALU (matched compliment) 226-461, >LINE2
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .473
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-DU1-aay-c-24-0-UI"
                     /tissue_type="Primary Lung Epithelial Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-DU1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-DU1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Epithelial Cells The
                     library was constructed according to Bonaldo, Lennon and
                     Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site.  Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
                     used to prime the synthesis of first-strand cDNA contains
                     a library tag sequence that is located between the Not I
                     site and the (dT)18 tail. The sequence tag for this
                     library is GGCTGTAGGC.
                     TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
                     TAG_LIB=UI-CF-DU1
                     TAG_SEQ=GGCTGTAGGC"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 3;  Length 473;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 19
```

```
RESULT 140
BM314257
LOCUS       BM314257                483 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  ig52c03.x1 HR85 islet Homo sapiens cDNA 3', mRNA sequence.
ACCESSION   BM314257
VERSION     BM314257.1  GI:18048602
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 3;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||
Db        377 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 420
```

```
RESULT 141
AQ276534/c
LOCUS       AQ276534                484 bp    DNA     linear   GSS 22-NOV-1998
DEFINITION  CITBI-E1-2521J15.TR CITBI-E1 Homo sapiens genomic clone 2521J15,
            genomic survey sequence.
ACCESSION   AQ276534
VERSION     AQ276534.1  GI:3902730
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .484
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2521J15"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 484;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
                ||||||||||||||||||||||||||||||||||||||||||||
Db        267 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 224


RESULT 142
BQ432539
LOCUS       BQ432539                485 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AGENCOURT_7905807 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6104737
            5', mRNA sequence.
ACCESSION   BQ432539
VERSION     BQ432539.1  GI:21171615
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2343  row: b  column: 02
            High quality sequence stop: 463.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6104737"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 485;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||
Db        371 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 414


RESULT 143
N35896
LOCUS       N35896                  486 bp    mRNA    linear   EST 16-JAN-1996
DEFINITION  yy28c08.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:272558 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N35896
VERSION     N35896.1  GI:1157038
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 440
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3882200"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:272558"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
```

/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||
Db 380 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 423

RESULT 144
AI908381
LOCUS AI908381 494 bp mRNA linear EST 30-MAR-2000
DEFINITION RC-BT170-280399-015 BT170 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI908381
VERSION AI908381.1 GI:6499061
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 494)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?t1=RC&t2=RC-BT170-015.html&t3=280399&t4=1)
Seq primer: puc 18 forward.
FEATURES Location/Qualifiers
source 1. .494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="Adult"
/clone_lib="BT170"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 494;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||
Db 378 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 421

RESULT 145
R86218/c
LOCUS R86218 495 bp mRNA linear EST 14-AUG-1995
DEFINITION yp88a10.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:194490 5' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION R86218
VERSION R86218.1 GI:944624
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 495)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 963
High quality sequence stops: 379
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 963 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 379.
FEATURES Location/Qualifiers
source 1. .495
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3763540"
/db_xref="taxon:9606"
/clone="IMAGE:194490"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

```
                    ||||||||||||||||||||||||||||||||||||||||||||
Db        255 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 212


RESULT 146
BE146359
LOCUS       BE146359                512 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  MR0-HT0209-010500-110-e11 HT0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE146359
VERSION     BE146359.1  GI:8609083
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 512)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR0-HT0209-010
            500-110-e11&t3=2000-05-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 126
            High quality sequence stop: 512.
FEATURES             Location/Qualifiers
     source          1. .512
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0209"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 512;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2979 CCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 3022
              ||||||||||||||||||||||||||||||||||||||||||||
Db        400 CCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 443


RESULT 147
AI888752
LOCUS       AI888752                516 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  wn36a07.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2447508 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI888752
VERSION     AI888752.1  GI:5593916
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1881   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2447508"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Gas4"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 516;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        395 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 438


RESULT 148
BE940436/c
LOCUS       BE940436                517 bp    mRNA    linear   EST 02-OCT-2000
DEFINITION  RC3-UT0062-210800-021-c05 UT0062 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE940436
VERSION     BE940436.1  GI:10469931
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 517)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
```

```
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-UT0062-210
            800-021-c05&t3=2000-08-21&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 9
            High quality sequence stop: 517.
FEATURES             Location/Qualifiers
     source          1. .517
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0062"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 2;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        354 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 311


RESULT 149
AQ026637
LOCUS       AQ026637                527 bp    DNA     linear   GSS 30-JUN-1998
DEFINITION  CIT-HSP-2314O17.TR CIT-HSP Homo sapiens genomic clone 2314O17,
            genomic survey sequence.
ACCESSION   AQ026637
VERSION     AQ026637.1  GI:3266859
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2314O17.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .527
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2314O17"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 527;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
              |||||||||||||||||||||||||||||||||||||||||||||
Db        215 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 258


RESULT 150
CN389220/c
LOCUS       CN389220                530 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600337068 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN389220
VERSION     CN389220.1  GI:47376815
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 530   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        416 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 373
```

```
RESULT 151
CD691041
LOCUS       CD691041                538 bp    mRNA    linear   EST 25-JUN-2003
DEFINITION  EST7564 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD691041
VERSION     CD691041.1  GI:32212370
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
. ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
  TITLE     Transcriptional Gene Expression Profile of Human Nasopharynx
  JOURNAL   Unpublished (2003)
COMMENT     Contact: YiXin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /note="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 538;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||
Db        326 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 369


RESULT 152
CV418654/c
LOCUS       CV418654                540 bp    mRNA    linear   EST 28-SEP-2004
DEFINITION  RC3-UT0062-210800-021-b11 UT0062 Homo sapiens cDNA, mRNA sequence.
ACCESSION   CV418654
VERSION     CV418654.1  GI:52814157
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 540)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. http://www.ludwig.org.br.
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0062"
                     /note="Organ: uterus tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||
Db        367 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 324


RESULT 153
B55060/c
LOCUS       B55060                  543 bp    DNA     linear   GSS 20-JUN-1998
DEFINITION  CIT-HSP-385P14.TF CIT-HSP Homo sapiens genomic clone 385P14,
            genomic survey sequence.
ACCESSION   B55060
VERSION     B55060.1  GI:2609394
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-385P14.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:5378919"
                     /db_xref="taxon:9606"
                     /clone="385P14"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
```

HindIII"
ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||
Db 436 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 393

RESULT 154
AQ713671/c
LOCUS AQ713671 543 bp DNA linear GSS 13-JUL-1999
DEFINITION HS_5399_B1_D03_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=975 Col=5 Row=H, genomic survey sequence.
ACCESSION AQ713671
VERSION AQ713671.1 GI:5462987
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 975 row: H column: 5
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 543.
FEATURES Location/Qualifiers
source 1. .543
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=975 Col=5 Row=H"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||
Db 391 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 348

RESULT 155
AA715817/c
LOCUS AA715817 551 bp mRNA linear EST 22-JAN-1998
DEFINITION nw25c05.s1 NCI_CGAP_GCB0 Homo sapiens cDNA clone IMAGE:1241480 3' similar to gb:K01144 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, GAMMA CHAIN PRECURSOR (HUMAN);contains Alu repetitive element;contains element PTR5 repetitive element ;, mRNA sequence.
ACCESSION AA715817
VERSION AA715817.1 GI:2728091
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1083 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 352.
FEATURES Location/Qualifiers
source 1. .551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1241480"
/tissue_type="germinal center B-cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_GCB0"
/note="Organ: tonsil; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Germinal center B-cells Library constructed by Dr. L. Staudt (NCI). 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||
Db 248 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 205

RESULT 156
CB129216/c
LOCUS CB129216 558 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0178834 L15CKK1 Homo sapiens cDNA clone L15CKK1-1-B09 5', mRNA sequence.
ACCESSION CB129216
VERSION CB129216.1 GI:28092697
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 1  row: B  column: 09
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L15CKK1-1-B09"
                     /sex="M"
                     /cell_line="CK-K1"
                     /lab_host="Top10F'"
                     /clone_lib="L15CKK1"
                     /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 558;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        310 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 267


RESULT 157
CV703188
LOCUS       CV703188                558 bp    mRNA    linear   EST 03-NOV-2004
DEFINITION  TgESTzys36b01.y1 Tg COUG Tachyzoite cDNA Library Toxoplasma gondii
            cDNA clone TgESTzys36b01.y1 5' similar to SW:ALU7_HUMAN P39194 ALU
            SUBFAMILY SQ SEQUENCE CONTAMINATION WARNING ENTRY. [1] ;, mRNA
            sequence.
ACCESSION   CV703188
VERSION     CV703188.1  GI:55285556
KEYWORDS    EST.
SOURCE      Toxoplasma gondii
  ORGANISM  Toxoplasma gondii
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioka,J.A., White,M.,
            Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
            Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
            Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I.,
            Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
  TITLE     Toxoplasma EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Clifton, S.
            Toxoplasma EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: toxo@watson.wustl.edu
            Contact David Sibley (toxoest@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Seq primer: T7 from Gibco
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Toxoplasma gondii"
                     /mol_type="mRNA"
                     /db_xref="taxon:5811"
                     /clone="TgESTzys36b01.y1"
                     /dev_stage="Tachyzoite"
                     /lab_host="GC10 Competent Cells(PGC)"
                     /clone_lib="Tg COUG Tachyzoite cDNA Library"
                     /note="Vector: pDNR-LIB; Site_1: SfiI; Site_2: SfiI; The
                     cDNA library was constructed by Keliang Tang, and Robert
                     Cole at Washington University. Total RNA was converted to
                     cDNA using the template-switching PCR method (Creator
                     SMART cDNA, Clontech Inc.). First strand was PCR amplified
                     using the same primer set and the fragments were digested
                     with SfiI. The fragments were size selected, ligated into
                     vector pDNR_LIB containing directional SfiI sites, and
                     electroporated into GC10 competent cells. WARNING: the
                     library contains a small percentage of cDNAs derived from
                     the human host cells."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 558;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        349 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 392


RESULT 158
AW877914
LOCUS       AW877914                563 bp    mRNA    linear   EST 23-MAY-2000
DEFINITION  MR3-OT0005-170200-101-c01 OT0005 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW877914
VERSION     AW877914.1  GI:8039833
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
```

```
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-OT0005-170
            200-101-c01&t3=2000-02-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 337.
FEATURES              Location/Qualifiers
     source           1. .563
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="OT0005"
                      /note="Organ: ovary; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 563;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db        263 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 306


RESULT 159
AQ591303
LOCUS       AQ591303                 569 bp    DNA     linear   GSS 08-JUN-1999
DEFINITION  HS_2122_B1_G04_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2122 Col=7 Row=N, genomic survey
            sequence.
ACCESSION   AQ591303
VERSION     AQ591303.1  GI:5022955
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 569)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2122  row: N  column: 7
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 569.
FEATURES              Location/Qualifiers
     source           1. .569
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=2122 Col=7 Row=N"
                      /sex="male"
                      /clone_lib="CIT Approved Human Genomic Sperm Library D"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 569;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
              |||||||||||||||||||||||||||||||||||||||||||||
Db        266 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 309


RESULT 160
CD723228/c
LOCUS       CD723228                 581 bp    mRNA    linear   EST 26-JUN-2003
DEFINITION  oj19e04.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
            clone oj19e04 5', mRNA sequence.
ACCESSION   CD723228
VERSION     CD723228.1  GI:32274076
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Ozyildirim,A.M., Wistow,G.J., Gao,J., Wang,J., Dickinson,D.P.,
            Frierson,H.F. Jr and Laurie,G.W.
  TITLE     The lacrimal gland transcriptome is an unusually rich source of
            rare and poorly characterized gene transcripts
  JOURNAL   Invest. Ophthalmol. Vis. Sci. 46 (5), 1572-1580 (2005)
   PUBMED   15851553
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 19  row: e  column: 04
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES              Location/Qualifiers
     source           1. .581
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="oj19e04"
                      /tissue_type="lacrimal gland"
                      /dev_stage="Adult"
                      /lab_host="EMDH10B"
                      /clone_lib="Human lacrimal gland, unamplified: oj"
                      /note="Organ: Eye; Vector: pCMVSport6; RNA was extracted
                      from 2 human lacrimal glands. A directionally cloned cDNA
                      library in the pCMVSPORT6 vector(Life Technologies) was
                      constructed at Bioserve Biotechnology (Laurel MD)
                      essentially following the protocols of the SuperScript
                      Plasmid System full details of which are contained in the
                      manufacturer's Instruction manual
                      (http://www.lifetech.com/). First strand synthesis was
                      carried out using a Not I primer-adapter
                      [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis
                      was performed on the unamplified library at the NIH
                      Intramural Sequencing Center (NISC)."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 581;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||
Db        138 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 95


RESULT 161
CB218356/c
LOCUS       CB218356                582 bp    mRNA    linear   EST 06-FEB-2003
DEFINITION  NISC_nb08c01.x1 COGENE 6E MAN Homo sapiens cDNA clone IMAGE:5795329
            3', mRNA sequence.
ACCESSION   CB218356
VERSION     CB218356.1  GI:28266548
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM12898  row: F  column: 2
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5795329"
                     /tissue_type="mandible, pooled"
                     /dev_stage="embryo, 6 weeks postconception"
                     /lab_host="DH10B"
                     /clone_lib="COGENE 6E MAN"
                     /note="Vector: pAMP1; cDNA primed using oligo-dT primer,
                     directionally cloned into UDG sites of pAMP1. Size
                     selected for insert sizes ranging from 0.2-2.0 kb.
                     Normalized to Cot5. Primary library, non-amplified.
                     Library constructed by M. Lovett. For more information on
                     this library, please contact R. Tidwell (Washington
                     University) or visit the COGENE website at
                     http://hg.wustl.edu/COGENE/."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 582;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
                ||||||||||||||||||||||||||||||||||||||||||||
Db         93 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 50


RESULT 162
AA206418/c
LOCUS       AA206418                590 bp    mRNA    linear   EST 12-MAR-1998
DEFINITION  zq51a05.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
            clone IMAGE:645104 3' similar to contains Alu repetitive
            element;contains element MER6 repetitive element ;, mRNA sequence.
ACCESSION   AA206418
VERSION     AA206418.1  GI:1801856
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1252   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 319.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:5215355"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:645104"
                     /dev_stage="Ntera-2/RA neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene neuroepithelium (#937231)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. NT2
                     cells (Ntera-2/c1.D1) induced with Retinoic Acid for 24
                     hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 590;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
                ||||||||||||||||||||||||||||||||||||||||||||
Db         68 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 25


RESULT 163
B70450
LOCUS       B70450                  602 bp    DNA     linear   GSS 21-JUN-1998
DEFINITION  CIT-HSP-2059C13.TF CIT-HSP Homo sapiens genomic clone 2059C13,
            genomic survey sequence.
ACCESSION   B70450
VERSION     B70450.1  GI:2709674
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
```

```
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7059904"
                     /db_xref="taxon:9606"
                     /clone="2059C13"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 602;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db         11 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 54


RESULT 164
AL691931
LOCUS       AL691931                614 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp313B1436_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313B1436 5', mRNA sequence.
ACCESSION   AL691931
VERSION     AL691931.1  GI:19617508
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp313B1436) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313B1436"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="313 (synonym: hlcc2)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        304 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 347


RESULT 165
CL246302
LOCUS       CL246302                614 bp    DNA     linear   GSS 22-JAN-2004
DEFINITION  HSC 01025 RPCI-11 Human Male BAC Library Homo sapiens genomic clone
            RP11-441N19, genomic survey sequence.
ACCESSION   CL246302
VERSION     CL246302.1  GI:41102792
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Scherer,S.W., Rommens,J.M. and Tsui,L.C.
  TITLE     Gene Identification on Human Chromosome 7 (Jan 05/04)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Scherer, S.W.; Rommens, J.M.; Tsui, L.C.
            The Human Chromosome 7 Project
            Department of Genetics, The Hospital for Sick Children
            555 University Avenue, Toronto, Ontario M5G 1X8, Canada
            Tel: 416 813 7613
            Fax: 416 813 8319
            Email: steve@genet.sickkids.on.ca
            Maps to human chromosome 7q11.23
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 614.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-441N19"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3057 GAGGTTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGG 3100
              ||||||||||||||||||||||||||||||||||||||||||||
Db        546 GAGGTTGCAATGAGCAAGATTGTGCCACTGCACTCCAGCCTGGG 589


RESULT 166
AL120523/c
LOCUS       AL120523                621 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp761O0048_s1 761 (synonym: hamy2) Homo sapiens cDNA clone
            DKFZp761O0048 3', mRNA sequence.
ACCESSION   AL120523
VERSION     AL120523.1  GI:5926422
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Koehrer, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            r1 sequence also available.
            This clone (DKFZp761O048) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp761O048"
                     /tissue_type="amygdala"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="761 (synonym: hamy2)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 621;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGAC 3114
              ||||||||||||||||||||||||||||||||||||||||||||
Db        151 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGAC 108


RESULT 167
CN267712
LOCUS       CN267712                627 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000531627977 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN267712
VERSION     CN267712.1  GI:47284126
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 627   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 7;  Length 627;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        583 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 626


RESULT 168
AQ542374/c
LOCUS       AQ542374                633 bp    DNA     linear   GSS 19-MAY-1999
DEFINITION  RPCI-11-346I12.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-346I12, genomic survey sequence.
ACCESSION   AQ542374
VERSION     AQ542374.1  GI:4872830
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-346I12.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7632683"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-346I12"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        507 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 464


RESULT 169
BX482192
LOCUS       BX482192                637 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686H14229_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686H14229 5', mRNA sequence.
ACCESSION   BX482192
VERSION     BX482192.1  GI:31941851
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686H14229) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686H14229"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 637;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db        389 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 432


RESULT 170
AQ697816/c
LOCUS       AQ697816                638 bp    DNA     linear   GSS 06-JUL-1999
DEFINITION  HS_5536_B1_B10_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1112 Col=19 Row=D, genomic survey sequence.
ACCESSION   AQ697816
VERSION     AQ697816.1  GI:5388064
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  PUBMED    10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1112  row: D  column: 19
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 638.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=1112 Col=19 Row=D"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 638;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db        353 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 310


RESULT 171
AV733586
LOCUS       AV733586                641 bp    mRNA    linear   EST 17-OCT-2000
DEFINITION  AV733586 cdA Homo sapiens cDNA clone cdAACG10 5', mRNA sequence.
ACCESSION   AV733586
VERSION     AV733586.1  GI:10851131
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S.,
            Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
            Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA cdA clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
```

```
     source          1. .641
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cdAACG10"
                     /tissue_type="pheochromocytoma"
                     /dev_stage="Adult"
                     /lab_host="BM25.8"
                     /clone_lib="cdA"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 641;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 449


RESULT 172
AG025905/c
LOCUS       AG025905                649 bp    DNA     linear   GSS 16-FEB-2005
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: B335N5 A043(Fw),
            genomic survey sequence.
ACCESSION   AG025905
VERSION     AG025905.1  GI:6594369
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in Database (1999)
REFERENCE   2  (bases 1 to 649)
  AUTHORS   Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
            Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1999) Masahira Hattori, RIKEN Genomic Sciences
            Center, RIKEN Yokohama Institute, Yokohama Research Promotion
            Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
            230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9113)
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q"
                     /clone="B335N5 A043(Fw)"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 649;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
Db        389 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 346


RESULT 173
AG080818
LOCUS       AG080818                650 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-077B13.R, genomic survey sequence.
ACCESSION   AG080818
VERSION     AG080818.1  GI:16632620
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 650)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-077B13.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 650;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 264


RESULT 174
AV722027
LOCUS       AV722027                652 bp    mRNA    linear   EST 16-OCT-2000
DEFINITION  AV722027 HTB Homo sapiens cDNA clone HTBBZD11 5', mRNA sequence.
ACCESSION   AV722027
VERSION     AV722027.1  GI:10824105
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
            Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L.,
            Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
            Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA HTB clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
```

```
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES              Location/Qualifiers
     source           1. .652
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="HTBBZD11"
                      /tissue_type="Hypothalamus"
                      /dev_stage="Adult"
                      /lab_host="SOLR"
                      /clone_lib="HTB"
                      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                      XhoI"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 652;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         25 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 68


RESULT 175
AG106557/c
LOCUS       AG106557                654 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-111B19.F, genomic survey sequence.
ACCESSION   AG106557
VERSION     AG106557.1  GI:16727075
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 654)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: -21M13
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES              Location/Qualifiers
     source           1. .654
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      /clone="PTB-111B19.F"
                      /sex="male"
                      /cell_type="lymphoblast"
                      /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        370 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 327


RESULT 176
AW959999
LOCUS       AW959999                655 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST372070 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW959999
VERSION     AW959999.1  GI:8149683
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 140
            Seq primer: Reverse.
FEATURES              Location/Qualifiers
     source           1. .655
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone_lib="MAGE resequences, MAGF"
                      /note="Vector: pBluescriptSKm"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        187 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 230


RESULT 177
AQ740363
LOCUS       AQ740363                694 bp    DNA     linear   GSS 16-JUL-1999
DEFINITION  HS_5501_A1_G11_SP6 RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1077 Col=21 Row=M, genomic survey sequence.
ACCESSION   AQ740363
VERSION     AQ740363.1  GI:5517885
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
```

```
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1077  row: M  column: 21
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=1077 Col=21 Row=M"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        156 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 199


RESULT 178
BU616112/c
LOCUS       BU616112                696 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  UI-H-DF0-bex-m-18-0-UI.s1 NCI_CGAP_DF0 Homo sapiens cDNA clone
            UI-H-DF0-bex-m-18-0-UI 3', mRNA sequence.
ACCESSION   BU616112
VERSION     BU616112.1  GI:23282320
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 696)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 11-299, >ALU (matched compliment) 647-696, >ALU
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .696
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DF0-bex-m-18-0-UI"
                     /tissue_type="Subchondral Bone"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DF0"
                     /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DF0 is a cDNA library containing the following
                     tissue(s): Subchondral Bone. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GTTAAGCGTC.
                     TAG_TISSUE=subchondral bone
                     TAG_LIB=UI-H-DF0
                     TAG_SEQ=GTTAAGCGTC"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 696;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||
Db         68 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 25


RESULT 179
AQ262202
LOCUS       AQ262202                706 bp    DNA     linear   GSS 24-OCT-1998
DEFINITION  CITBI-E1-2509J18.TF CITBI-E1 Homo sapiens genomic clone 2509J18,
            genomic survey sequence.
ACCESSION   AQ262202
VERSION     AQ262202.1  GI:3788682
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 706)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2509J18.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .706
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:9606"
                     /clone="2509J18"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 706;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        538 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 581


RESULT 180
BX504396
LOCUS       BX504396                715 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686O22133_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686O22133 5', mRNA sequence.
ACCESSION   BX504396
VERSION     BX504396.1  GI:32031347
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686O22133) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686O22133"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3072 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||
Db        255 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 298


RESULT 181
BX098422/c
LOCUS       BX098422                716 bp    mRNA    linear   EST 04-FEB-2003
DEFINITION  BX098422 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGp998J0883 ; IMAGE:109927, mRNA sequence.
ACCESSION   BX098422
VERSION     BX098422.1  GI:27829399
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998J0883.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998J0883 ; IMAGE:109927"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 5;  Length 716;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        140 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 97


RESULT 182
AG175100/c
LOCUS       AG175100                745 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-045O24.TJ, genomic survey
            sequence.
ACCESSION   AG175100
VERSION     AG175100.1  GI:16704780
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
```

```
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 745)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                Sequencing: TJ
             LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES              Location/Qualifiers
     source           1. .745
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      /clone="RP43-045O24.TJ"
                      /sex="male"
                      /cell_type="lymphocytes"
                      /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 10;  Length 745;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
              ||||||||||||||||||||||||||||||||||||||||||||
Db        478 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 435


RESULT 183
CB308313
LOCUS       CB308313                781 bp    mRNA    linear   EST 04-MAR-2003
DEFINITION  AGENCOURT_11818308 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6882307 5', mRNA sequence.
ACCESSION   CB308313
VERSION     CB308313.1  GI:28831023
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
             cDNA Library Preparation: CLONTECH
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3128  row: 1  column: 18
            High quality sequence stop: 544.
FEATURES              Location/Qualifiers
     source           1. .781
                      /organism="Macaca mulatta"
                      /mol_type="mRNA"
                      /db_xref="taxon:9544"
                      /clone="IMAGE:6882307"
                      /tissue_type="Ovary"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NICHD_Rh_Ov1"
                      /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
                      Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                      Average insert size 1.0-4.0 kb. Tissue pooled from
                      pre-pubertal, post pubertal sn menopausal monkeys.
                      Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 781;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db        205 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 248


RESULT 184
BZ610701/c
LOCUS       BZ610701                803 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHAAJ70TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-3K20, genomic survey sequence.
ACCESSION   BZ610701
VERSION     BZ610701.1  GI:31519262
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .803
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="MCF7_1-3K20"
                      /sex="female"
                      /clone_lib="Human MCF7 breast cancer cell line library
                      (MCF7_1)"
                      /note="Vector: pECBAC1; Site_1: HindIII; This library was
                      constructed from MCF7 breast cancer cell line by Amplicon
                      Express (http://www.genomex.com) using their standard
                      procedure."
ORIGIN
```

```
  Query Match             1.4%;  Score 44;  DB 9;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 178


RESULT 185
CA442440
LOCUS       CA442440                828 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-DP0-ave-m-01-0-UI.s1 NCI_CGAP_Fs1 Homo sapiens cDNA clone
            UI-H-DP0-ave-m-01-0-UI 3', mRNA sequence.
ACCESSION   CA442440
VERSION     CA442440.1  GI:24806860
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 163-455, >ALU 416-466, >ALU
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .828
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-DP0-ave-m-01-0-UI"
                     /tissue_type="Fibrosarcoma"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Fs1"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_Fs1 is
                     a cDNA library containing the following tissue(s):
                     Fibrosarcoma Cell line HT-1088 (ATCC number CCL-121). The
                     library was constructed according to Bonaldo, Lennon and
                     Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site. Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
                     used to prime the synthesis of first-strand cDNA contains
                     a library tag sequence that is located between the Not I
                     site and the (dT)18 tail. The sequence tag for this
                     library is GTTCTACGAG.
                     TAG_TISSUE=fibrosarcoma
                     TAG_LIB=UI-H-DP0
                     TAG_SEQ=GTTCTACGAG"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 828;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        398 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 441


RESULT 186
CB307968/c
LOCUS       CB307968                830 bp    mRNA    linear   EST 04-MAR-2003
DEFINITION  AGENCOURT_11771960 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6882266 5', mRNA sequence.
ACCESSION   CB307968
VERSION     CB307968.1  GI:28830678
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
             cDNA Library Preparation: CLONTECH
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3128  row: k  column: 01
            High quality sequence stop: 499.
FEATURES             Location/Qualifiers
     source          1. .830
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:6882266"
                     /tissue_type="Ovary"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_Rh_Ov1"
                     /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
                     Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.0-4.0 kb. Tissue pooled from
                     pre-pubertal, post pubertal sn menopausal monkeys.
                     Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 6;  Length 830;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        365 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 322


RESULT 187
AQ781442/c
LOCUS       AQ781442                830 bp    DNA     linear   GSS 02-AUG-1999
DEFINITION  HS_3117_A2_G10_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3117 Col=20 Row=M, genomic survey
            sequence.
ACCESSION   AQ781442
VERSION     AQ781442.1  GI:5684402
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
```

```
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3117  row: M  column: 20
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 830.
FEATURES             Location/Qualifiers
     source          1. .830
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3117 Col=20 Row=M"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 830;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3072 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db        454 AAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 411


RESULT 188
BZ612098/c
LOCUS       BZ612098                834 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHADP72TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-22L24, genomic survey sequence.
ACCESSION   BZ612098
VERSION     BZ612098.1  GI:31520659
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     End-sequence profiling: Sequence-based analysis of aberrant genomes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
   PUBMED   12788976
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-22L24"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 834;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        204 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 161


RESULT 189
AQ781550/c
LOCUS       AQ781550                847 bp    DNA     linear   GSS 02-AUG-1999
DEFINITION  HS_3176_A2_A08_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3176 Col=16 Row=A, genomic survey
            sequence.
ACCESSION   AQ781550
VERSION     AQ781550.1  GI:5684510
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3176  row: A  column: 16
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 847.
FEATURES             Location/Qualifiers
     source          1. .847
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3176 Col=16 Row=A"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 9;  Length 847;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAAC 2938
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAAC 279
```

RESULT 190
BF573285
LOCUS BF573285 867 bp mRNA linear EST 12-DEC-2000
DEFINITION 602079678F2 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4254115 5', mRNA sequence.
ACCESSION BF573285
VERSION BF573285.1 GI:11646997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1081 row: 1 column: 20
High quality sequence stop: 290.
FEATURES Location/Qualifiers
source 1. .867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4254115"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_62"
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN

Query Match 1.4%; Score 44; DB 2; Length 867;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        228 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 271
```

RESULT 191
AQ749032/c
LOCUS AQ749032 903 bp DNA linear GSS 19-JUL-1999
DEFINITION HS_5575_A2_E04_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1151 Col=8 Row=I, genomic survey sequence.
ACCESSION AQ749032
VERSION AQ749032.1 GI:5536190
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 1151 row: I column: 8
Seq primer: T7
Class: BAC ends
High quality sequence stop: 903.
FEATURES Location/Qualifiers
source 1. .903
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1151 Col=8 Row=I"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 903;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||
Db        359 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 316
```

RESULT 192
AL570050
LOCUS AL570050 982 bp mRNA linear EST 05-APR-2004
DEFINITION AL570050 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI006YK15 3-PRIME, mRNA sequence.
ACCESSION AL570050
VERSION AL570050.3 GI:46236290
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31291480.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

```
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            6704.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI006AF08NP1&c=6704.f.
FEATURES              Location/Qualifiers
     source           1. .982
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DI006YK15"
                      /tissue_type="PLACENTA COT 25-NORMALIZED"
                      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             1.4%;  Score 44;  DB 1;  Length 982;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 394


RESULT 193
CR622010/c
LOCUS       CR622010                1527 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI006YK15 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR622010
VERSION     CR622010.1  GI:50502817
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1527)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1527)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES              Location/Qualifiers
     source           1. .1527
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DI006YK15"
                      /tissue_type="Placenta Cot 25-normalized"
                      /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             1.4%;  Score 44;  DB 4;  Length 1527;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-09;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db       1296 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 1253


RESULT 194
BE141837/c
LOCUS       BE141837                108 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  IL5-HT0117-170999-004-A02 HT0117 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE141837
VERSION     BE141837.1  GI:8604558
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL5-HT0117-170
            999-004-A02&t3=1999-09-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 108.
FEATURES              Location/Qualifiers
     source           1. .108
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="HT0117"
                      /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 43;  DB 2;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 6e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3080 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db         60 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 18


RESULT 195
AA728876
```

```
LOCUS       AA728876                132 bp    mRNA    linear   EST 22-JAN-1998
DEFINITION  nv37c08.r1 NCI_CGAP_Br5 Homo sapiens cDNA clone IMAGE:1222382
            similar to contains Alu repetitive element;contains element MSR1
            repetitive element ;, mRNA sequence.
ACCESSION   AA728876
VERSION     AA728876.1  GI:2750235
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 132)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,
            Ph.D. student, Rodrigo F. Chuaqui, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 180   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .132
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1222382"
                     /sex="female"
                     /tissue_type="infiltrating ductal carcinoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Br5"
                     /note="Organ: breast; Vector: pAMP10; mRNA made from
                     infiltrating ductal carcinoma, cDNA made by oligo-dT
                     priming.  Non-directionally cloned.  Size-selected on
                     agarose  gel, average insert size 600 bp.  "
ORIGIN

  Query Match             1.4%;  Score 43;  DB 1;  Length 132;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||
Db         67 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 109


RESULT 196
BG012771/c
LOCUS       BG012771                132 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  IL5-GN0239-141200-341-d03 GN0239 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG012771
VERSION     BG012771.1  GI:12462298
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 132)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL5&t2=IL5-GN0239-
            141200-341-d03&t3=2000-12-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 14
            High quality sequence stop: 132.
FEATURES             Location/Qualifiers
     source          1. .132
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0239"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

  Query Match             1.4%;  Score 43;  DB 2;  Length 132;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||
Db        127 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85


RESULT 197
CR543084/c
LOCUS       CR543084                151 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  DKFZp470G0412_r1 470 (synonym: pliv1) Pongo pygmaeus cDNA clone
            DKFZp470G0412 5', mRNA sequence.
ACCESSION   CR543084
VERSION     CR543084.1  GI:49894295
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 151)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Wambutt,R., Heubner,D., Mewes,H.W., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa
            (Berlin/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp470G0412) is available at
            the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
            Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
```

```
            clone@rzpd.de Further information about the clone and the
            sequencing project is available at
            http://mips.gsf.de/projects/cdna/.
FEATURES              Location/Qualifiers
     source           1. .151
                      /organism="Pongo pygmaeus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9600"
                      /clone="DKFZp470G0412"
                      /tissue_type="liver"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="470 (synonym: pliv1)"
                      /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             1.4%;  Score 43;  DB 7;  Length 151;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3080 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db         63 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 21


RESULT 198
N73776/c
LOCUS       N73776                  154 bp    mRNA    linear   EST 19-MAR-1996
DEFINITION  za61h08.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:297087 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N73776
VERSION     N73776.1  GI:1231061
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 131.
FEATURES              Location/Qualifiers
     source           1. .154
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:1242008"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:297087"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 43;  DB 8;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3080 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db         43 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 1


RESULT 199
AI611895
LOCUS       AI611895                157 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION  tt59f11.x1 NCI_CGAP_HSC4 Homo sapiens cDNA clone IMAGE:2245101 3'
            similar to contains Alu repetitive element;contains element TAR1
            repetitive element ;, mRNA sequence.
ACCESSION   AI611895
VERSION     AI611895.1  GI:4621062
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 157)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 215   Std Error: 0.00
            Seq primer: -40UP from Gibco
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .157
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2245101"
                      /tissue_type="CD34+, CD38- from normal bone marrow donor"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_HSC4"
                      /note="Organ: bone marrow; Vector: pAMP1; mRNA made from
                      lymphoid tissue, cDNA made by oligo-dT  priming.
                      Directionally cloned.  Size-selected on agarose  gel,
                      average insert size 500 bp.  Primary library,
                      non-amplified.   cDNA Library Preparation: David B.
                      Krizman, Ph.D. Reference:  Krizman et al. (1996) Cancer
                      Research 56:5380-5383."
ORIGIN

  Query Match             1.4%;  Score 43;  DB 1;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3080 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||
Db         75 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 117
```

RESULT 200
AA018823/c
LOCUS AA018823 168 bp mRNA linear EST 30-JAN-1997
DEFINITION ze57e09.r1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:363112 5' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AA018823
VERSION AA018823.1 GI:1482215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1965 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 139.
FEATURES Location/Qualifiers
source 1. .168
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1279815"
/db_xref="taxon:9606"
/clone="IMAGE:363112"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares retina N2b4HR"
/note="Organ: eye; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo. "
ORIGIN

Query Match 1.4%; Score 43; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.9e-09;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3080 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||
Db 128 GCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 86

Search completed: May 11, 2006, 12:00:24
Job time : 11860 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:27:55 ; Search time 531 Seconds
(without alignments)
10451.125 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2605320

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	52	1.7	381	3	US-09-513-999C-3793	Sequence 3793, Ap
c	2	52	1.7	601	3	US-09-949-016-72610	Sequence 72610, A
	3	52	1.7	601	3	US-09-949-016-75710	Sequence 75710, A
	4	52	1.7	601	3	US-09-949-016-75723	Sequence 75723, A
c	5	52	1.7	15415	3	US-09-949-016-13931	Sequence 13931, A
c	6	52	1.7	15415	3	US-09-949-016-13932	Sequence 13932, A
c	7	52	1.7	25969	3	US-09-949-016-13397	Sequence 13397, A
c	8	52	1.7	77586	3	US-09-949-016-13220	Sequence 13220, A
c	9	52	1.7	77586	3	US-09-949-016-13221	Sequence 13221, A
	10	52	1.7	92139	3	US-09-918-686-1	Sequence 1, Appli
	11	52	1.7	120727	3	US-09-949-016-15787	Sequence 15787, A
	12	52	1.7	120727	3	US-09-949-016-15788	Sequence 15788, A
	13	52	1.7	155266	3	US-09-949-016-13870	Sequence 13870, A
c	14	51	1.6	35688	3	US-09-949-016-16873	Sequence 16873, A
	15	51	1.6	156942	3	US-09-949-016-12227	Sequence 12227, A
	16	51	1.6	156950	3	US-09-949-016-15946	Sequence 15946, A
c	17	50	1.6	255	3	US-09-513-999C-28539	Sequence 28539, A
	18	50	1.6	288	3	US-09-973-278-870	Sequence 870, App
	19	50	1.6	601	3	US-09-949-016-47563	Sequence 47563, A
	20	50	1.6	601	3	US-09-949-016-125579	Sequence 125579,
	21	50	1.6	601	3	US-09-949-016-125580	Sequence 125580,
	22	50	1.6	601	3	US-09-949-016-125581	Sequence 125581,
	23	50	1.6	601	3	US-09-949-016-163881	Sequence 163881,
	24	50	1.6	8705	3	US-09-544-398B-10	Sequence 10, Appl
	25	50	1.6	8705	3	US-09-543-771B-10	Sequence 10, Appl
	26	50	1.6	36954	3	US-09-949-016-15304	Sequence 15304, A
	27	50	1.6	39154	3	US-09-949-016-12384	Sequence 12384, A
	28	50	1.6	39154	3	US-09-949-016-12801	Sequence 12801, A
	29	50	1.6	39443	3	US-09-949-016-14326	Sequence 14326, A
	30	50	1.6	39443	3	US-09-949-016-14327	Sequence 14327, A
	31	50	1.6	76401	3	US-09-949-016-17153	Sequence 17153, A
	32	50	1.6	87190	3	US-09-949-016-16335	Sequence 16335, A
c	33	50	1.6	95648	3	US-09-949-016-13139	Sequence 13139, A
	34	50	1.6	140844	3	US-09-949-016-14199	Sequence 14199, A
	35	50	1.6	197336	3	US-09-949-016-12881	Sequence 12881, A
	36	50	1.6	197337	3	US-09-949-016-14376	Sequence 14376, A
	37	50	1.6	197337	3	US-09-949-002-738	Sequence 738, App
	38	50	1.6	234287	3	US-09-949-002-687	Sequence 687, App
	39	50	1.6	234288	3	US-09-949-016-17272	Sequence 17272, A
	40	50	1.6	234288	3	US-09-949-002-841	Sequence 841, App
c	41	50	1.6	254964	3	US-09-949-016-12583	Sequence 12583, A
c	42	50	1.6	254964	3	US-09-949-016-17392	Sequence 17392, A
c	43	47	1.5	601	3	US-09-949-016-164699	Sequence 164699,
	44	47	1.5	12880	3	US-09-949-016-16733	Sequence 16733, A
c	45	47	1.5	45833	3	US-09-949-016-14330	Sequence 14330, A
c	46	47	1.5	45833	3	US-09-949-016-14331	Sequence 14331, A
c	47	47	1.5	45833	3	US-09-949-016-14332	Sequence 14332, A
c	48	47	1.5	45833	3	US-09-949-016-14333	Sequence 14333, A
	49	47	1.5	67386	3	US-09-949-016-16519	Sequence 16519, A
c	50	47	1.5	108310	3	US-09-949-016-16366	Sequence 16366, A
	51	47	1.5	139049	3	US-09-949-016-17030	Sequence 17030, A
	52	46	1.5	601	3	US-09-949-016-65112	Sequence 65112, A
	53	46	1.5	601	3	US-09-949-016-65113	Sequence 65113, A
	54	46	1.5	601	3	US-09-949-016-65114	Sequence 65114, A
	55	46	1.5	601	3	US-09-949-016-137514	Sequence 137514,
c	56	46	1.5	601	3	US-09-949-016-155725	Sequence 155725,
c	57	46	1.5	601	3	US-09-949-016-155726	Sequence 155726,
c	58	46	1.5	601	3	US-09-949-016-169304	Sequence 169304,
c	59	46	1.5	601	3	US-09-949-016-169305	Sequence 169305,
c	60	46	1.5	601	3	US-09-949-016-169306	Sequence 169306,
	61	46	1.5	601	3	US-09-949-016-198268	Sequence 198268,
	62	46	1.5	601	3	US-09-949-016-198269	Sequence 198269,
	63	46	1.5	601	3	US-09-949-016-198270	Sequence 198270,
	64	46	1.5	601	3	US-09-949-016-206260	Sequence 206260,
	65	46	1.5	601	3	US-09-949-016-206261	Sequence 206261,
	66	46	1.5	601	3	US-09-949-016-206262	Sequence 206262,
	67	46	1.5	17704	3	US-09-949-016-14878	Sequence 14878, A
c	68	46	1.5	27600	3	US-09-949-016-15290	Sequence 15290, A
c	69	46	1.5	29960	3	US-09-949-016-11978	Sequence 11978, A
c	70	46	1.5	30054	3	US-09-949-016-16100	Sequence 16100, A
c	71	46	1.5	40905	3	US-09-949-016-16864	Sequence 16864, A
c	72	46	1.5	60917	3	US-09-949-016-12638	Sequence 12638, A
c	73	46	1.5	60917	3	US-09-949-016-16057	Sequence 16057, A
c	74	46	1.5	84916	3	US-09-949-016-14736	Sequence 14736, A
	75	46	1.5	96739	3	US-09-949-016-15606	Sequence 15606, A
c	76	46	1.5	116955	3	US-09-949-016-17565	Sequence 17565, A
c	77	46	1.5	118136	3	US-09-949-016-12439	Sequence 12439, A
c	78	46	1.5	144322	3	US-09-949-016-15316	Sequence 15316, A
c	79	46	1.5	177293	3	US-09-949-016-16513	Sequence 16513, A
	80	46	1.5	240157	3	US-09-949-016-16264	Sequence 16264, A
c	81	46	1.5	253345	3	US-09-949-016-12656	Sequence 12656, A
c	82	46	1.5	253364	3	US-09-949-016-13639	Sequence 13639, A
c	83	46	1.5	331814	3	US-09-949-016-12008	Sequence 12008, A
c	84	46	1.5	331814	3	US-09-949-016-17056	Sequence 17056, A
c	85	46	1.5	767677	3	US-09-949-016-12147	Sequence 12147, A
c	86	46	1.5	767677	3	US-09-949-016-17361	Sequence 17361, A
	87	45	1.4	478	3	US-09-621-976-15071	Sequence 15071, A
	88	45	1.4	601	3	US-09-949-016-30237	Sequence 30237, A
	89	45	1.4	601	3	US-09-949-016-30238	Sequence 30238, A
	90	45	1.4	601	3	US-09-949-016-30239	Sequence 30239, A
c	91	45	1.4	601	3	US-09-949-016-65461	Sequence 65461, A
	92	45	1.4	601	3	US-09-949-016-72556	Sequence 72556, A
	93	45	1.4	601	3	US-09-949-016-72557	Sequence 72557, A
	94	45	1.4	601	3	US-09-949-016-82095	Sequence 82095, A
	95	45	1.4	601	3	US-09-949-016-82096	Sequence 82096, A
	96	45	1.4	601	3	US-09-949-016-82097	Sequence 82097, A
c	97	45	1.4	601	3	US-09-949-016-145830	Sequence 145830,

c 98	45	1.4	601	3	US-09-949-016-146098	Sequence 146098,
c 99	45	1.4	601	3	US-09-949-016-146366	Sequence 146366,
100	45	1.4	601	3	US-09-949-016-166442	Sequence 166442,
c 101	45	1.4	601	3	US-09-949-016-169130	Sequence 169130,
c 102	45	1.4	601	3	US-09-949-016-169131	Sequence 169131,
103	45	1.4	601	3	US-09-949-016-203717	Sequence 203717,
104	45	1.4	601	3	US-09-949-016-203718	Sequence 203718,
105	45	1.4	601	3	US-09-949-016-203719	Sequence 203719,
106	45	1.4	601	3	US-09-949-002-4958	Sequence 4958, Ap
107	45	1.4	601	3	US-09-949-002-4959	Sequence 4959, Ap
c 108	45	1.4	601	3	US-09-949-002-9370	Sequence 9370, Ap
c 109	45	1.4	8429	3	US-09-949-016-14483	Sequence 14483, A
c 110	45	1.4	9123	3	US-09-949-016-14852	Sequence 14852, A
c 111	45	1.4	9176	3	US-09-949-016-17432	Sequence 17432, A
112	45	1.4	10597	3	US-09-949-016-13766	Sequence 13766, A
c 113	45	1.4	15975	3	US-09-949-016-17469	Sequence 17469, A
114	45	1.4	16595	3	US-09-146-053-7	Sequence 7, Appli
c 115	45	1.4	19296	3	US-09-949-016-16393	Sequence 16393, A
c 116	45	1.4	19943	3	US-09-949-016-15920	Sequence 15920, A
c 117	45	1.4	19945	3	US-09-949-016-12553	Sequence 12553, A
c 118	45	1.4	27545	3	US-09-949-016-13069	Sequence 13069, A
119	45	1.4	33099	3	US-09-949-016-16094	Sequence 16094, A
c 120	45	1.4	36223	3	US-09-949-016-14417	Sequence 14417, A
c 121	45	1.4	36546	3	US-09-949-016-17543	Sequence 17543, A
122	45	1.4	40090	3	US-09-820-004-3	Sequence 3, Appli
123	45	1.4	40147	3	US-09-949-016-12374	Sequence 12374, A
124	45	1.4	40379	3	US-09-949-016-14128	Sequence 14128, A
125	45	1.4	40877	3	US-09-949-016-17495	Sequence 17495, A
126	45	1.4	40908	3	US-09-949-002-845	Sequence 845, App
127	45	1.4	40944	3	US-09-949-002-603	Sequence 603, App
128	45	1.4	47781	3	US-09-949-016-16492	Sequence 16492, A
129	45	1.4	47781	3	US-09-949-016-16493	Sequence 16493, A
130	45	1.4	47781	3	US-09-949-016-16494	Sequence 16494, A
c 131	45	1.4	49677	3	US-09-949-002-817	Sequence 817, App
132	45	1.4	64984	3	US-09-949-016-15254	Sequence 15254, A
c 133	45	1.4	66627	3	US-09-949-016-12630	Sequence 12630, A
c 134	45	1.4	66628	3	US-09-949-016-16112	Sequence 16112, A
135	45	1.4	70313	3	US-09-949-002-714	Sequence 714, App
136	45	1.4	72347	3	US-09-949-002-633	Sequence 633, App
c 137	45	1.4	80858	3	US-09-949-016-12659	Sequence 12659, A
c 138	45	1.4	80859	3	US-09-949-016-15715	Sequence 15715, A
139	45	1.4	87648	3	US-09-949-016-13655	Sequence 13655, A
140	45	1.4	91279	3	US-09-949-016-15146	Sequence 15146, A
c 141	45	1.4	93532	3	US-09-949-016-15944	Sequence 15944, A
142	45	1.4	94142	3	US-09-949-016-16553	Sequence 16553, A
143	45	1.4	96690	3	US-09-949-016-17103	Sequence 17103, A
c 144	45	1.4	103750	3	US-09-949-016-13319	Sequence 13319, A
145	45	1.4	112465	3	US-09-949-016-16424	Sequence 16424, A
146	45	1.4	120609	3	US-09-949-016-13915	Sequence 13915, A
c 147	45	1.4	140725	3	US-09-949-016-17074	Sequence 17074, A
c 148	45	1.4	155266	3	US-09-949-016-13870	Sequence 13870, A
149	45	1.4	162841	3	US-09-949-016-13733	Sequence 13733, A
c 150	45	1.4	174639	3	US-09-949-016-16509	Sequence 16509, A
151	45	1.4	205044	3	US-09-949-016-15851	Sequence 15851, A
152	45	1.4	205044	3	US-09-949-016-15852	Sequence 15852, A
153	45	1.4	205044	3	US-09-949-016-15853	Sequence 15853, A
c 154	45	1.4	218940	3	US-09-949-016-17539	Sequence 17539, A
155	45	1.4	223471	3	US-09-949-016-12387	Sequence 12387, A
156	45	1.4	223471	3	US-09-949-016-12724	Sequence 12724, A
157	45	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A
c 158	44	1.4	145	3	US-09-513-999C-25910	Sequence 25910, A
159	44	1.4	330	3	US-09-513-999C-28305	Sequence 28305, A
160	44	1.4	601	3	US-09-949-016-33113	Sequence 33113, A
161	44	1.4	601	3	US-09-949-016-36581	Sequence 36581, A
162	44	1.4	601	3	US-09-949-016-38281	Sequence 38281, A
163	44	1.4	601	3	US-09-949-016-46100	Sequence 46100, A
164	44	1.4	601	3	US-09-949-016-46119	Sequence 46119, A
165	44	1.4	601	3	US-09-949-016-71064	Sequence 71064, A
166	44	1.4	601	3	US-09-949-016-71065	Sequence 71065, A
167	44	1.4	601	3	US-09-949-016-71066	Sequence 71066, A
168	44	1.4	601	3	US-09-949-016-71067	Sequence 71067, A
169	44	1.4	601	3	US-09-949-016-71068	Sequence 71068, A
170	44	1.4	601	3	US-09-949-016-71069	Sequence 71069, A
c 171	44	1.4	601	3	US-09-949-016-121487	Sequence 121487,
c 172	44	1.4	601	3	US-09-949-016-121488	Sequence 121488,
c 173	44	1.4	601	3	US-09-949-016-133933	Sequence 133933,
c 174	44	1.4	601	3	US-09-949-016-133934	Sequence 133934,
175	44	1.4	601	3	US-09-949-016-143264	Sequence 143264,
176	44	1.4	601	3	US-09-949-016-143435	Sequence 143435,
c 177	44	1.4	601	3	US-09-949-016-144443	Sequence 144443,
c 178	44	1.4	601	3	US-09-949-016-144444	Sequence 144444,
179	44	1.4	601	3	US-09-949-016-150959	Sequence 150959,
180	44	1.4	601	3	US-09-949-016-163060	Sequence 163060,
c 181	44	1.4	601	3	US-09-949-016-165548	Sequence 165548,
c 182	44	1.4	601	3	US-09-949-016-199110	Sequence 199110,
183	44	1.4	601	3	US-09-949-016-202729	Sequence 202729,
184	44	1.4	601	3	US-09-949-016-204472	Sequence 204472,
185	44	1.4	601	3	US-09-949-016-204542	Sequence 204542,
c 186	44	1.4	601	3	US-09-949-002-10588	Sequence 10588, A
187	44	1.4	956	3	US-09-641-638-56	Sequence 56, Appl
188	44	1.4	956	3	US-10-170-097-56	Sequence 56, Appl
189	44	1.4	7480	3	US-09-949-016-11999	Sequence 11999, A
190	44	1.4	7567	3	US-09-949-016-16486	Sequence 16486, A
c 191	44	1.4	7720	3	US-09-318-448-5	Sequence 5, Appli
c 192	44	1.4	7720	3	US-09-577-266-5	Sequence 5, Appli
c 193	44	1.4	10720	3	US-09-949-016-17443	Sequence 17443, A
194	44	1.4	15116	3	US-09-949-016-12354	Sequence 12354, A
195	44	1.4	15117	3	US-09-949-016-16260	Sequence 16260, A
196	44	1.4	18651	3	US-09-949-002-592	Sequence 592, App
197	44	1.4	18682	3	US-09-949-002-786	Sequence 786, App
c 198	44	1.4	18891	3	US-09-949-016-16523	Sequence 16523, A
199	44	1.4	19389	3	US-09-949-016-15113	Sequence 15113, A
200	44	1.4	19974	3	US-09-949-016-15267	Sequence 15267, A
201	44	1.4	20674	3	US-09-641-638-651	Sequence 651, App
202	44	1.4	20674	3	US-10-170-097-651	Sequence 651, App
c 203	44	1.4	22906	3	US-09-949-016-16471	Sequence 16471, A
c 204	44	1.4	25160	3	US-09-949-016-12691	Sequence 12691, A
c 205	44	1.4	25162	3	US-09-949-016-13065	Sequence 13065, A
c 206	44	1.4	25162	3	US-09-949-016-13066	Sequence 13066, A
c 207	44	1.4	26007	3	US-09-949-016-17222	Sequence 17222, A
c 208	44	1.4	28809	3	US-09-949-016-12825	Sequence 12825, A
209	44	1.4	29930	3	US-09-949-016-15326	Sequence 15326, A
c 210	44	1.4	31444	3	US-09-949-016-17532	Sequence 17532, A
c 211	44	1.4	33870	3	US-09-949-016-16306	Sequence 16306, A
c 212	44	1.4	34310	3	US-09-949-016-12518	Sequence 12518, A
c 213	44	1.4	34315	3	US-09-949-016-15469	Sequence 15469, A
214	44	1.4	36618	3	US-09-949-016-12961	Sequence 12961, A
c 215	44	1.4	38503	3	US-09-949-016-17257	Sequence 17257, A
216	44	1.4	38969	3	US-09-949-016-14051	Sequence 14051, A
c 217	44	1.4	39715	3	US-09-949-016-12454	Sequence 12454, A
c 218	44	1.4	48691	3	US-09-949-016-16308	Sequence 16308, A
c 219	44	1.4	54986	3	US-09-949-016-16716	Sequence 16716, A
c 220	44	1.4	58782	3	US-09-949-016-16851	Sequence 16851, A
221	44	1.4	60137	3	US-09-949-016-14735	Sequence 14735, A
222	44	1.4	60137	3	US-09-949-016-14912	Sequence 14912, A
c 223	44	1.4	64291	3	US-09-949-016-16278	Sequence 16278, A
224	44	1.4	64319	3	US-09-949-016-12804	Sequence 12804, A
225	44	1.4	65848	3	US-09-949-016-13285	Sequence 13285, A
226	44	1.4	68778	3	US-09-949-016-16406	Sequence 16406, A
227	44	1.4	78491	3	US-09-949-016-15132	Sequence 15132, A
228	44	1.4	82000	3	US-09-949-016-15595	Sequence 15595, A
c 229	44	1.4	85850	3	US-09-949-016-13424	Sequence 13424, A
c 230	44	1.4	85869	3	US-09-949-016-12017	Sequence 12017, A
c 231	44	1.4	85878	3	US-09-949-016-16321	Sequence 16321, A
c 232	44	1.4	88268	3	US-09-949-016-14178	Sequence 14178, A
233	44	1.4	89584	3	US-09-949-016-17068	Sequence 17068, A
234	44	1.4	92681	3	US-09-949-016-14772	Sequence 14772, A
c 235	44	1.4	93510	3	US-09-949-016-15095	Sequence 15095, A
c 236	44	1.4	97195	3	US-09-949-016-12212	Sequence 12212, A
c 237	44	1.4	97196	3	US-09-949-016-16971	Sequence 16971, A
238	44	1.4	105001	3	US-09-949-002-627	Sequence 627, App
239	44	1.4	105002	3	US-09-949-002-821	Sequence 821, App
c 240	44	1.4	107941	3	US-09-949-016-14206	Sequence 14206, A
241	44	1.4	109038	3	US-09-949-016-12199	Sequence 12199, A
c 242	44	1.4	109378	3	US-09-949-016-12391	Sequence 12391, A
243	44	1.4	112874	3	US-09-949-016-13180	Sequence 13180, A

244	44	1.4	114183	3	US-09-949-002-849	Sequence 849, App
c 245	44	1.4	117410	3	US-09-949-016-12262	Sequence 12262, A
c 246	44	1.4	120727	3	US-09-949-016-15787	Sequence 15787, A
c 247	44	1.4	120727	3	US-09-949-016-15788	Sequence 15788, A
248	44	1.4	127546	3	US-09-949-002-624	Sequence 624, App
249	44	1.4	133613	3	US-09-949-016-15824	Sequence 15824, A
c 250	44	1.4	134008	3	US-09-949-016-13841	Sequence 13841, A
c 251	44	1.4	134434	3	US-09-949-016-17362	Sequence 17362, A
c 252	44	1.4	139257	3	US-09-920-671-11	Sequence 11, Appl
c 253	44	1.4	139936	3	US-09-949-016-11782	Sequence 11782, A
c 254	44	1.4	139952	3	US-09-949-016-13280	Sequence 13280, A
255	44	1.4	146307	3	US-09-949-016-14881	Sequence 14881, A
256	44	1.4	146307	3	US-09-949-016-14882	Sequence 14882, A
257	44	1.4	146307	3	US-09-949-016-14883	Sequence 14883, A
258	44	1.4	146307	3	US-09-949-016-14884	Sequence 14884, A
259	44	1.4	146307	3	US-09-949-016-14885	Sequence 14885, A
260	44	1.4	146307	3	US-09-949-016-14886	Sequence 14886, A
261	44	1.4	146307	3	US-09-949-016-14887	Sequence 14887, A
262	44	1.4	146307	3	US-09-949-016-14888	Sequence 14888, A
263	44	1.4	148405	3	US-09-949-016-11747	Sequence 11747, A
264	44	1.4	148405	3	US-09-949-016-12835	Sequence 12835, A
265	44	1.4	148405	3	US-09-949-016-12836	Sequence 12836, A
266	44	1.4	148405	3	US-09-949-016-12837	Sequence 12837, A
267	44	1.4	152481	3	US-09-949-016-12521	Sequence 12521, A
268	44	1.4	152798	3	US-09-949-016-12775	Sequence 12775, A
269	44	1.4	152822	3	US-09-949-016-17518	Sequence 17518, A
270	44	1.4	152822	3	US-09-949-016-17519	Sequence 17519, A
c 271	44	1.4	157822	3	US-09-949-016-16723	Sequence 16723, A
c 272	44	1.4	205044	3	US-09-949-016-15851	Sequence 15851, A
c 273	44	1.4	205044	3	US-09-949-016-15852	Sequence 15852, A
c 274	44	1.4	205044	3	US-09-949-016-15853	Sequence 15853, A
c 275	44	1.4	211049	3	US-09-949-016-15770	Sequence 15770, A
c 276	44	1.4	221958	3	US-09-949-016-12173	Sequence 12173, A
c 277	44	1.4	221966	3	US-09-949-016-15498	Sequence 15498, A
c 278	44	1.4	223471	3	US-09-949-016-12387	Sequence 12387, A
c 279	44	1.4	223471	3	US-09-949-016-12724	Sequence 12724, A
c 280	44	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A
281	44	1.4	237510	3	US-09-949-016-14273	Sequence 14273, A
282	44	1.4	239527	3	US-09-949-016-15980	Sequence 15980, A
283	44	1.4	250352	3	US-09-949-016-14724	Sequence 14724, A
284	44	1.4	275110	3	US-09-949-016-12706	Sequence 12706, A
285	44	1.4	275110	3	US-09-949-016-16070	Sequence 16070, A
286	44	1.4	304533	3	US-09-949-016-15371	Sequence 15371, A
287	44	1.4	304533	3	US-09-949-016-15372	Sequence 15372, A
288	44	1.4	385136	3	US-09-949-016-16073	Sequence 16073, A
289	44	1.4	393753	3	US-09-949-016-14573	Sequence 14573, A
290	44	1.4	393753	3	US-09-949-016-14574	Sequence 14574, A
291	44	1.4	784019	3	US-09-949-016-14033	Sequence 14033, A
292	44	1.4	818128	3	US-09-949-016-14546	Sequence 14546, A
293	44	1.4	818128	3	US-09-949-016-14547	Sequence 14547, A
294	44	1.4	818128	3	US-09-949-016-14548	Sequence 14548, A
295	44	1.4	818128	3	US-09-949-016-14549	Sequence 14549, A
296	44	1.4	818128	3	US-09-949-016-14550	Sequence 14550, A
297	44	1.4	818128	3	US-09-949-016-14551	Sequence 14551, A
298	44	1.4	818128	3	US-09-949-016-14552	Sequence 14552, A
299	44	1.4	818128	3	US-09-949-016-14553	Sequence 14553, A
300	44	1.4	818128	3	US-09-949-016-14554	Sequence 14554, A
301	44	1.4	818128	3	US-09-949-016-14555	Sequence 14555, A
302	44	1.4	818128	3	US-09-949-016-14556	Sequence 14556, A
303	44	1.4	818128	3	US-09-949-016-14557	Sequence 14557, A
304	44	1.4	818128	3	US-09-949-016-14558	Sequence 14558, A
305	44	1.4	818128	3	US-09-949-016-14559	Sequence 14559, A
306	44	1.4	818128	3	US-09-949-016-14560	Sequence 14560, A
307	44	1.4	818128	3	US-09-949-016-14561	Sequence 14561, A
308	44	1.4	818128	3	US-09-949-016-14562	Sequence 14562, A
309	44	1.4	818128	3	US-09-949-016-14564	Sequence 14564, A
310	44	1.4	818128	3	US-09-949-016-14565	Sequence 14565, A
311	44	1.4	818128	3	US-09-949-016-14566	Sequence 14566, A
312	44	1.4	818128	3	US-09-949-016-14567	Sequence 14567, A
313	44	1.4	828152	3	US-09-949-016-12777	Sequence 12777, A
c 314	43	1.4	117	3	US-09-513-999C-17724	Sequence 17724, A
315	43	1.4	354	3	US-09-973-278-811	Sequence 811, App
316	43	1.4	356	3	US-09-973-278-804	Sequence 804, App
c 317	43	1.4	365	3	US-09-513-999C-21034	Sequence 21034, A
318	43	1.4	601	3	US-09-949-016-17714	Sequence 17714, A
319	43	1.4	601	3	US-09-949-016-20309	Sequence 20309, A
320	43	1.4	601	3	US-09-949-016-20310	Sequence 20310, A
321	43	1.4	601	3	US-09-949-016-21707	Sequence 21707, A
322	43	1.4	601	3	US-09-949-016-21708	Sequence 21708, A
323	43	1.4	601	3	US-09-949-016-21709	Sequence 21709, A
c 324	43	1.4	601	3	US-09-949-016-26397	Sequence 26397, A
325	43	1.4	601	3	US-09-949-016-35107	Sequence 35107, A
326	43	1.4	601	3	US-09-949-016-35108	Sequence 35108, A
327	43	1.4	601	3	US-09-949-016-45237	Sequence 45237, A
328	43	1.4	601	3	US-09-949-016-48203	Sequence 48203, A
c 329	43	1.4	601	3	US-09-949-016-56368	Sequence 56368, A
c 330	43	1.4	601	3	US-09-949-016-56369	Sequence 56369, A
331	43	1.4	601	3	US-09-949-016-66550	Sequence 66550, A
c 332	43	1.4	601	3	US-09-949-016-89155	Sequence 89155, A
c 333	43	1.4	601	3	US-09-949-016-103330	Sequence 103330,
c 334	43	1.4	601	3	US-09-949-016-103331	Sequence 103331,
335	43	1.4	601	3	US-09-949-016-105600	Sequence 105600,
336	43	1.4	601	3	US-09-949-016-132803	Sequence 132803,
337	43	1.4	601	3	US-09-949-016-132804	Sequence 132804,
c 338	43	1.4	601	3	US-09-949-016-137031	Sequence 137031,
c 339	43	1.4	601	3	US-09-949-016-137032	Sequence 137032,
c 340	43	1.4	601	3	US-09-949-016-138675	Sequence 138675,
341	43	1.4	601	3	US-09-949-016-142312	Sequence 142312,
342	43	1.4	601	3	US-09-949-016-142313	Sequence 142313,
c 343	43	1.4	601	3	US-09-949-016-144926	Sequence 144926,
344	43	1.4	601	3	US-09-949-016-167908	Sequence 167908,
345	43	1.4	601	3	US-09-949-016-167909	Sequence 167909,
346	43	1.4	601	3	US-09-949-016-175295	Sequence 175295,
347	43	1.4	601	3	US-09-949-016-179586	Sequence 179586,
348	43	1.4	601	3	US-09-949-016-179587	Sequence 179587,
349	43	1.4	601	3	US-09-949-016-179588	Sequence 179588,
c 350	43	1.4	601	3	US-09-949-016-183547	Sequence 183547,
c 351	43	1.4	601	3	US-09-949-016-196942	Sequence 196942,
352	43	1.4	5232	3	US-09-212-971-3	Sequence 3, Appli
353	43	1.4	5232	3	US-08-800-929A-3	Sequence 3, Appli
354	43	1.4	5232	3	US-09-617-053A-3	Sequence 3, Appli
c 355	43	1.4	5372	3	US-09-949-016-14484	Sequence 14484, A
c 356	43	1.4	5386	3	US-09-949-016-11885	Sequence 11885, A
357	43	1.4	5749	3	US-09-949-016-15441	Sequence 15441, A
358	43	1.4	5749	3	US-09-949-016-15442	Sequence 15442, A
c 359	43	1.4	10013	3	US-09-949-016-16474	Sequence 16474, A
c 360	43	1.4	15033	3	US-09-949-016-15916	Sequence 15916, A
361	43	1.4	20902	3	US-09-949-016-12684	Sequence 12684, A
362	43	1.4	20908	3	US-09-949-016-14764	Sequence 14764, A
363	43	1.4	23946	3	US-09-949-016-13381	Sequence 13381, A
c 364	43	1.4	25370	3	US-09-949-016-12109	Sequence 12109, A
c 365	43	1.4	25375	3	US-09-949-016-15880	Sequence 15880, A
c 366	43	1.4	31229	3	US-09-949-016-12619	Sequence 12619, A
c 367	43	1.4	31231	3	US-09-949-016-15766	Sequence 15766, A
368	43	1.4	32584	3	US-09-949-016-16766	Sequence 16766, A
369	43	1.4	33152	3	US-09-949-016-16719	Sequence 16719, A
c 370	43	1.4	36816	3	US-09-949-016-16218	Sequence 16218, A
c 371	43	1.4	39318	3	US-09-949-016-13798	Sequence 13798, A
c 372	43	1.4	42246	3	US-09-949-016-17008	Sequence 17008, A
c 373	43	1.4	42574	3	US-09-949-016-17525	Sequence 17525, A
374	43	1.4	45427	3	US-09-949-016-16243	Sequence 16243, A
c 375	43	1.4	48763	3	US-09-916-204-3	Sequence 3, Appli
c 376	43	1.4	48763	3	US-10-282-048-3	Sequence 3, Appli
c 377	43	1.4	49309	3	US-09-949-016-12680	Sequence 12680, A
c 378	43	1.4	49309	3	US-09-949-016-13084	Sequence 13084, A
379	43	1.4	50775	3	US-09-949-016-12858	Sequence 12858, A
380	43	1.4	50776	3	US-09-949-016-15438	Sequence 15438, A
381	43	1.4	51508	3	US-09-949-016-16681	Sequence 16681, A
382	43	1.4	57914	3	US-09-949-016-11935	Sequence 11935, A
383	43	1.4	57936	3	US-09-949-016-16921	Sequence 16921, A
c 384	43	1.4	58014	3	US-09-949-016-17448	Sequence 17448, A
c 385	43	1.4	59123	3	US-09-949-016-12177	Sequence 12177, A
386	43	1.4	59977	3	US-09-949-016-12259	Sequence 12259, A
387	43	1.4	59978	3	US-09-949-016-15397	Sequence 15397, A
c 388	43	1.4	62386	3	US-09-949-016-12823	Sequence 12823, A
389	43	1.4	75378	3	US-09-949-016-17140	Sequence 17140, A

c 390	43	1.4	76118	3	US-09-949-016-15593	Sequence 15593, A
c 391	43	1.4	76269	3	US-09-949-016-14603	Sequence 14603, A
c 392	43	1.4	76553	3	US-09-949-016-13432	Sequence 13432, A
393	43	1.4	83210	3	US-09-949-016-14209	Sequence 14209, A
394	43	1.4	87752	3	US-09-949-016-16807	Sequence 16807, A
395	43	1.4	90618	3	US-09-949-016-15964	Sequence 15964, A
c 396	43	1.4	91559	3	US-09-949-016-12581	Sequence 12581, A
c 397	43	1.4	91559	3	US-09-949-016-13701	Sequence 13701, A
398	43	1.4	93364	3	US-09-949-016-14890	Sequence 14890, A
399	43	1.4	93894	3	US-09-949-016-13629	Sequence 13629, A
400	43	1.4	99797	3	US-09-949-016-15255	Sequence 15255, A
c 401	43	1.4	106256	3	US-09-949-016-16858	Sequence 16858, A
402	43	1.4	109250	3	US-09-949-016-12530	Sequence 12530, A
403	43	1.4	109251	3	US-09-949-016-17321	Sequence 17321, A
404	43	1.4	113379	3	US-09-949-016-17561	Sequence 17561, A
c 405	43	1.4	113379	3	US-09-949-016-17561	Sequence 17561, A
406	43	1.4	113379	3	US-09-949-016-17562	Sequence 17562, A
c 407	43	1.4	113379	3	US-09-949-016-17562	Sequence 17562, A
c 408	43	1.4	120213	3	US-09-949-016-13304	Sequence 13304, A
c 409	43	1.4	120217	3	US-09-949-016-12260	Sequence 12260, A
c 410	43	1.4	122772	3	US-09-949-016-14132	Sequence 14132, A
411	43	1.4	129899	3	US-09-949-016-14684	Sequence 14684, A
c 412	43	1.4	144596	3	US-09-949-016-11749	Sequence 11749, A
c 413	43	1.4	144596	3	US-09-949-016-13035	Sequence 13035, A
414	43	1.4	148156	3	US-09-949-016-11776	Sequence 11776, A
c 415	43	1.4	154626	3	US-09-949-016-14000	Sequence 14000, A
c 416	43	1.4	160018	3	US-09-949-016-12617	Sequence 12617, A
c 417	43	1.4	160018	3	US-09-949-016-15994	Sequence 15994, A
418	43	1.4	160552	3	US-09-593-828-11	Sequence 11, Appl
c 419	43	1.4	173992	3	US-09-949-016-13379	Sequence 13379, A
c 420	43	1.4	174493	3	US-09-804-471A-3	Sequence 3, Appli
c 421	43	1.4	174493	3	US-10-238-709-3	Sequence 3, Appli
c 422	43	1.4	174493	3	US-10-724-594-3	Sequence 3, Appli
c 423	43	1.4	177251	3	US-09-949-016-15841	Sequence 15841, A
c 424	43	1.4	194714	3	US-09-949-016-11869	Sequence 11869, A
c 425	43	1.4	194889	3	US-09-949-016-15654	Sequence 15654, A
c 426	43	1.4	196714	3	US-09-949-016-15474	Sequence 15474, A
c 427	43	1.4	237510	3	US-09-949-016-14273	Sequence 14273, A
c 428	43	1.4	346112	3	US-09-949-016-13165	Sequence 13165, A
429	42	1.3	336	3	US-09-385-982-17	Sequence 17, Appl
430	42	1.3	601	3	US-09-949-016-19618	Sequence 19618, A
c 431	42	1.3	601	3	US-09-949-016-25518	Sequence 25518, A
432	42	1.3	601	3	US-09-949-016-28910	Sequence 28910, A
433	42	1.3	601	3	US-09-949-016-31419	Sequence 31419, A
434	42	1.3	601	3	US-09-949-016-37684	Sequence 37684, A
435	42	1.3	601	3	US-09-949-016-37704	Sequence 37704, A
c 436	42	1.3	601	3	US-09-949-016-39927	Sequence 39927, A
c 437	42	1.3	601	3	US-09-949-016-45798	Sequence 45798, A
438	42	1.3	601	3	US-09-949-016-54520	Sequence 54520, A
c 439	42	1.3	601	3	US-09-949-016-57030	Sequence 57030, A
c 440	42	1.3	601	3	US-09-949-016-57821	Sequence 57821, A
c 441	42	1.3	601	3	US-09-949-016-57822	Sequence 57822, A
442	42	1.3	601	3	US-09-949-016-59688	Sequence 59688, A
c 443	42	1.3	601	3	US-09-949-016-73696	Sequence 73696, A
444	42	1.3	601	3	US-09-949-016-86369	Sequence 86369, A
445	42	1.3	601	3	US-09-949-016-86370	Sequence 86370, A
446	42	1.3	601	3	US-09-949-016-105326	Sequence 105326,
c 447	42	1.3	601	3	US-09-949-016-121736	Sequence 121736,
448	42	1.3	601	3	US-09-949-016-125857	Sequence 125857,
449	42	1.3	601	3	US-09-949-016-125964	Sequence 125964,
450	42	1.3	601	3	US-09-949-016-127231	Sequence 127231,
c 451	42	1.3	601	3	US-09-949-016-127826	Sequence 127826,
c 452	42	1.3	601	3	US-09-949-016-128163	Sequence 128163,
453	42	1.3	601	3	US-09-949-016-135139	Sequence 135139,
c 454	42	1.3	601	3	US-09-949-016-139090	Sequence 139090,
455	42	1.3	601	3	US-09-949-016-141088	Sequence 141088,
c 456	42	1.3	601	3	US-09-949-016-146571	Sequence 146571,
457	42	1.3	601	3	US-09-949-016-147357	Sequence 147357,
c 458	42	1.3	601	3	US-09-949-016-169230	Sequence 169230,
459	42	1.3	601	3	US-09-949-016-171811	Sequence 171811,
460	42	1.3	601	3	US-09-949-016-171812	Sequence 171812,
461	42	1.3	601	3	US-09-949-016-175700	Sequence 175700,
c 462	42	1.3	601	3	US-09-949-016-176743	Sequence 176743,
c 463	42	1.3	601	3	US-09-949-016-177625	Sequence 177625,
c 464	42	1.3	601	3	US-09-949-016-180839	Sequence 180839,
c 465	42	1.3	601	3	US-09-949-016-186854	Sequence 186854,
c 466	42	1.3	601	3	US-09-949-016-188045	Sequence 188045,
c 467	42	1.3	601	3	US-09-949-016-188074	Sequence 188074,
468	42	1.3	601	3	US-09-949-016-188551	Sequence 188551,
c 469	42	1.3	601	3	US-09-949-016-199254	Sequence 199254,
c 470	42	1.3	601	3	US-09-949-016-199255	Sequence 199255,
471	42	1.3	601	3	US-09-949-002-1276	Sequence 1276, Ap
472	42	1.3	601	3	US-09-949-002-1277	Sequence 1277, Ap
473	42	1.3	601	3	US-09-949-002-10408	Sequence 10408, A
474	42	1.3	601	3	US-09-949-002-10409	Sequence 10409, A
c 475	42	1.3	5227	3	US-09-973-278-919	Sequence 919, App
c 476	42	1.3	5235	3	US-09-973-278-918	Sequence 918, App
c 477	42	1.3	5969	3	US-09-949-016-1342	Sequence 1342, Ap
c 478	42	1.3	5970	3	US-09-949-016-938	Sequence 938, App
c 479	42	1.3	6328	3	US-09-949-016-14847	Sequence 14847, A
c 480	42	1.3	6831	3	US-09-949-016-12281	Sequence 12281, A
c 481	42	1.3	7584	3	US-09-949-016-11836	Sequence 11836, A
c 482	42	1.3	7585	3	US-09-949-016-16737	Sequence 16737, A
483	42	1.3	9900	3	US-09-949-016-13392	Sequence 13392, A
c 484	42	1.3	11022	3	US-09-949-016-15502	Sequence 15502, A
c 485	42	1.3	15661	3	US-09-949-016-13161	Sequence 13161, A
c 486	42	1.3	15664	3	US-09-949-016-14031	Sequence 14031, A
c 487	42	1.3	18264	3	US-09-949-016-13030	Sequence 13030, A
488	42	1.3	18605	3	US-09-949-016-14465	Sequence 14465, A
489	42	1.3	20059	3	US-09-949-016-12323	Sequence 12323, A
490	42	1.3	20060	3	US-09-949-016-14716	Sequence 14716, A
491	42	1.3	20084	3	US-09-949-016-14165	Sequence 14165, A
c 492	42	1.3	23995	3	US-09-949-016-17363	Sequence 17363, A
493	42	1.3	24841	3	US-09-949-016-14009	Sequence 14009, A
c 494	42	1.3	24847	3	US-09-949-016-16056	Sequence 16056, A
495	42	1.3	27592	3	US-09-949-016-14682	Sequence 14682, A
496	42	1.3	27767	3	US-09-949-016-13055	Sequence 13055, A
497	42	1.3	27923	3	US-09-949-016-13371	Sequence 13371, A
c 498	42	1.3	28791	3	US-09-949-016-15396	Sequence 15396, A
499	42	1.3	28823	3	US-09-949-016-12437	Sequence 12437, A
500	42	1.3	28823	3	US-09-949-016-15879	Sequence 15879, A

ALIGNMENTS

```
RESULT 1
US-09-513-999C-3793/c
; Sequence 3793, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 3793
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 12..347
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 12..158
;   OTHER INFORMATION: score 7.7
;   OTHER INFORMATION: seq LCLLGSSDSPASA/SQ
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: 322
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 370
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 376
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 55
;   OTHER INFORMATION: Xaa=Ser or Thr
US-09-513-999C-3793

  Query Match             1.7%;  Score 52;  DB 3;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 56


RESULT 2
US-09-949-016-72610/c
; Sequence 72610, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72610
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-72610

  Query Match             1.7%;  Score 52;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 20


RESULT 3
US-09-949-016-75710
; Sequence 75710, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75710
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-75710

  Query Match             1.7%;  Score 52;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 242


RESULT 4
US-09-949-016-75723
; Sequence 75723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75723
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-75723

  Query Match             1.7%;  Score 52;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 242


RESULT 5
US-09-949-016-13931/c
; Sequence 13931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13931
;   LENGTH: 15415
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13931

  Query Match             1.7%;  Score 52;  DB 3;  Length 15415;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11938 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 11887


RESULT 6
US-09-949-016-13932/c
; Sequence 13932, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13932
;   LENGTH: 15415
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13932

  Query Match             1.7%;  Score 52;  DB 3;  Length 15415;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11938 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 11887


RESULT 7
US-09-949-016-13397/c
; Sequence 13397, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13397
;   LENGTH: 25969
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13397

  Query Match             1.7%;  Score 52;  DB 3;  Length 25969;
  Best Local Similarity   100.0%;  Pred. No. 3e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22148 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 22097


RESULT 8
US-09-949-016-13220/c
; Sequence 13220, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13220
;   LENGTH: 77586
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13220

  Query Match             1.7%;  Score 52;  DB 3;  Length 77586;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 43027


RESULT 9
US-09-949-016-13221/c
; Sequence 13221, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13221
;   LENGTH: 77586
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13221

  Query Match             1.7%;  Score 52;  DB 3;  Length 77586;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-14;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43078 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 43027
```

RESULT 10
US-09-918-686-1
; Sequence 1, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paeper, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043, 8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match 1.7%; Score 52; DB 3; Length 92139;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85294 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85345
```

RESULT 11
US-09-949-016-15787
; Sequence 15787, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15787
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15787

Query Match 1.7%; Score 52; DB 3; Length 120727;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58522 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 58573
```

RESULT 12
US-09-949-016-15788
; Sequence 15788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15788
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15788

Query Match 1.7%; Score 52; DB 3; Length 120727;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58522 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 58573
```

RESULT 13
US-09-949-016-13870
; Sequence 13870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13870
; LENGTH: 155266
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13870

Query Match 1.7%; Score 52; DB 3; Length 155266;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     139149 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 139200


RESULT 14
US-09-949-016-16873/c
; Sequence 16873, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16873
;   LENGTH: 35688
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16873

  Query Match             1.6%;  Score 51;  DB 3;  Length 35688;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-14;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7571 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 7521


RESULT 15
US-09-949-016-12227
; Sequence 12227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12227
;   LENGTH: 156942
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(156942)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

  Query Match             1.6%;  Score 51;  DB 3;  Length 156942;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-14;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118483 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 118533


RESULT 16
US-09-949-016-15946
; Sequence 15946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15946
;   LENGTH: 156950
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(156950)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15946

  Query Match             1.6%;  Score 51;  DB 3;  Length 156950;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-14;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118483 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 118533


RESULT 17
US-09-513-999C-28539/c
; Sequence 28539, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 28539
;   LENGTH: 255
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-28539

  Query Match             1.6%;  Score 50;  DB 3;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 177
```

```
RESULT 18
US-09-973-278-870
; Sequence 870, Application US/09973278
; Patent No. 6924354
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P2
;  CURRENT APPLICATION NUMBER: US/09/973,278
;  CURRENT FILING DATE:  2001-10-10
;  PRIOR APPLICATION NUMBER: 60/239,899
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 09/227,357
;  PRIOR FILING DATE: 1999-01-08
;  PRIOR APPLICATION NUMBER: PCT/US98/13684
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/051,926
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,793
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,925
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,929
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,803
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,732
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,931
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,932
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,916
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,930
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,918
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,920
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,733
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,795
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,919
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,928
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/055,722
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,723
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,948
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,949
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,953
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,950
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,947
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,964
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/056,360
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,684
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,984
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,954
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/058,785
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,664
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,660
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,661
;  PRIOR FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 947
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 870
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-973-278-870

  Query Match             1.6%;  Score 50;  DB 3;  Length 288;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 270


RESULT 19
US-09-949-016-47563
; Sequence 47563, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47563
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-47563

  Query Match             1.6%;  Score 50;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 580


RESULT 20
US-09-949-016-125579
; Sequence 125579, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
```

```
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125579
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-125579

  Query Match             1.6%;  Score 50;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 270


RESULT 21
US-09-949-016-125580
; Sequence 125580, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125580
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-125580

  Query Match             1.6%;  Score 50;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 252


RESULT 22
US-09-949-016-125581
; Sequence 125581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125581
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-125581

  Query Match             1.6%;  Score 50;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          46 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 95


RESULT 23
US-09-949-016-163881
; Sequence 163881, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163881
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-163881

  Query Match             1.6%;  Score 50;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 228


RESULT 24
US-09-544-398B-10
; Sequence 10, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
;  APPLICANT: Carulli, John P.
;  APPLICANT:  Little, Randall D.
;  APPLICANT:  Recker, Robert R.
;  APPLICANT:  Johnson, Mark L.
;  TITLE OF INVENTION: High bone mass gene of 11q13.3
;  FILE REFERENCE: 032796-013
;  CURRENT APPLICATION NUMBER: US/09/544,398B
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/229,319
;  PRIOR FILING DATE: 1999-01-13
;  PRIOR APPLICATION NUMBER: US 60/071,449
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/105,511
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 641
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-544-398B-10

  Query Match             1.6%;  Score 50;  DB 3;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 25
US-09-543-771B-10
; Sequence 10, Application US/09543771B
; Patent No. 6780609
; GENERAL INFORMATION:
;  APPLICANT: Carulli, John P.
;  APPLICANT:  Little, Randall D.
;  APPLICANT:  Recker, Robert R.
;  APPLICANT:  Johnson, Mark L.
;  TITLE OF INVENTION: High bone mass gene of 11q13.3
;  FILE REFERENCE: 032796-014
;  CURRENT APPLICATION NUMBER: US/09/543,771B
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 09/229,319
;  PRIOR FILING DATE: 1999-01-13
;  PRIOR APPLICATION NUMBER: US 60/071,449
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/105,511
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 641
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-543-771B-10

  Query Match             1.6%;  Score 50;  DB 3;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 26
US-09-949-016-15304
; Sequence 15304, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15304
;   LENGTH: 36954
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15304

  Query Match             1.6%;  Score 50;  DB 3;  Length 36954;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27720 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 27769


RESULT 27
US-09-949-016-12384
; Sequence 12384, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12384
;   LENGTH: 39154
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39154)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12384

  Query Match             1.6%;  Score 50;  DB 3;  Length 39154;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38617 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38666


RESULT 28
US-09-949-016-12801
; Sequence 12801, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12801
;   LENGTH: 39154
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39154)
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12801

  Query Match             1.6%;  Score 50;  DB 3;  Length 39154;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38617 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38666


RESULT 29
US-09-949-016-14326
; Sequence 14326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14326
;   LENGTH: 39443
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39443)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14326

  Query Match             1.6%;  Score 50;  DB 3;  Length 39443;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38617 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38666


RESULT 30
US-09-949-016-14327
; Sequence 14327, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14327
;   LENGTH: 39443
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39443)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14327

  Query Match             1.6%;  Score 50;  DB 3;  Length 39443;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38617 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 38666


RESULT 31
US-09-949-016-17153
; Sequence 17153, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17153
;   LENGTH: 76401
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17153

  Query Match             1.6%;  Score 50;  DB 3;  Length 76401;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56563 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 56612


RESULT 32
US-09-949-016-16335
; Sequence 16335, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16335
;   LENGTH: 87190
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(87190)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16335

  Query Match             1.6%;  Score 50;  DB 3;  Length 87190;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57590 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 57639


RESULT 33
US-09-949-016-13139/c
; Sequence 13139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13139
;   LENGTH: 95648
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(95648)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13139

  Query Match             1.6%;  Score 50;  DB 3;  Length 95648;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68111 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 68062


RESULT 34
US-09-949-016-14199
; Sequence 14199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14199
;   LENGTH: 140844
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(140844)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14199

  Query Match             1.6%;  Score 50;  DB 3;  Length 140844;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140232 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 140281


RESULT 35
US-09-949-016-12881
; Sequence 12881, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12881
;   LENGTH: 197336
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(197336)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12881

  Query Match             1.6%;  Score 50;  DB 3;  Length 197336;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 36
US-09-949-016-14376
; Sequence 14376, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14376
;   LENGTH: 197337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(197337)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14376

  Query Match             1.6%;  Score 50;  DB 3;  Length 197337;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 37
US-09-949-002-738
; Sequence 738, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 738
;   LENGTH: 197337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(197337)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-002-738

  Query Match             1.6%;  Score 50;  DB 3;  Length 197337;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 38
US-09-949-002-687
; Sequence 687, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 687
;   LENGTH: 234287
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(234287)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-002-687

  Query Match             1.6%;  Score 50;  DB 3;  Length 234287;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 39
US-09-949-016-17272
; Sequence 17272, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17272
;   LENGTH: 234288
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(234288)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17272

  Query Match             1.6%;  Score 50;  DB 3;  Length 234288;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 40
US-09-949-002-841
; Sequence 841, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 841
;   LENGTH: 234288
```

```
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(234288)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-002-841

  Query Match             1.6%;  Score 50;  DB 3;  Length 234288;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 189483


RESULT 41
US-09-949-016-12583/c
; Sequence 12583, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12583
;   LENGTH: 254964
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(254964)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12583

  Query Match             1.6%;  Score 50;  DB 3;  Length 254964;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108004 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 107955


RESULT 42
US-09-949-016-17392/c
; Sequence 17392, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17392
;   LENGTH: 254964
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(254964)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17392

  Query Match             1.6%;  Score 50;  DB 3;  Length 254964;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108004 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 107955


RESULT 43
US-09-949-016-164699/c
; Sequence 164699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 164699
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-164699

  Query Match             1.5%;  Score 47;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-12;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 3119
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        580 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 534


RESULT 44
US-09-949-016-16733
; Sequence 16733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 16733
; LENGTH: 12880
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16733

Query Match 1.5%; Score 47; DB 3; Length 12880;
Best Local Similarity 100.0%; Pred. No. 6.8e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGC 2934
||||||||||||||||||||||||||||||||||||||||||||||||
Db 5157 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGC 5203

RESULT 45
US-09-949-016-14330/c
; Sequence 14330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14330
; LENGTH: 45833
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14330

Query Match 1.5%; Score 47; DB 3; Length 45833;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
||||||||||||||||||||||||||||||||||||||||||||||||
Db 22968 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 22922

RESULT 46
US-09-949-016-14331/c
; Sequence 14331, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14331
; LENGTH: 45833
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14331

Query Match 1.5%; Score 47; DB 3; Length 45833;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
||||||||||||||||||||||||||||||||||||||||||||||||
Db 22968 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 22922

RESULT 47
US-09-949-016-14332/c
; Sequence 14332, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14332
; LENGTH: 45833
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14332

Query Match 1.5%; Score 47; DB 3; Length 45833;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
||||||||||||||||||||||||||||||||||||||||||||||||
Db 22968 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 22922

RESULT 48
US-09-949-016-14333/c
; Sequence 14333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14333
; LENGTH: 45833
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14333

Query Match 1.5%; Score 47; DB 3; Length 45833;
Best Local Similarity 100.0%; Pred. No. 6.5e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22968 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 22922


RESULT 49
US-09-949-016-16519
; Sequence 16519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16519
;   LENGTH: 67386
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(67386)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16519

  Query Match             1.5%;  Score 47;  DB 3;  Length 67386;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-12;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      49989 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 50035


RESULT 50
US-09-949-016-16366/c
; Sequence 16366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16366
;   LENGTH: 108310
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(108310)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16366

  Query Match             1.5%;  Score 47;  DB 3;  Length 108310;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-12;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 3119
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      63285 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 63239


RESULT 51
US-09-949-016-17030
; Sequence 17030, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17030
;   LENGTH: 139049
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(139049)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17030

  Query Match             1.5%;  Score 47;  DB 3;  Length 139049;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-12;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 3119
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      42519 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGT 42565


RESULT 52
US-09-949-016-65112
; Sequence 65112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65112
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-65112

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
```

```
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 112


RESULT 53
US-09-949-016-65113
; Sequence 65113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65113
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-65113

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 440


RESULT 54
US-09-949-016-65114
; Sequence 65114, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65114
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-65114

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 55
US-09-949-016-137514
; Sequence 137514, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 137514
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-137514

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        324 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 369


RESULT 56
US-09-949-016-155725/c
; Sequence 155725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155725
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-155725

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 441


RESULT 57
US-09-949-016-155726/c
; Sequence 155726, Application US/09949016
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155726
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-155726

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        477 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 432


RESULT 58
US-09-949-016-169304/c
; Sequence 169304, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169304
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169304

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        468 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 423


RESULT 59
US-09-949-016-169305/c
; Sequence 169305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169305
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169305

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        353 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 308


RESULT 60
US-09-949-016-169306/c
; Sequence 169306, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169306
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169306

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 166


RESULT 61
US-09-949-016-198268
; Sequence 198268, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
```

```
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198268
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198268

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 367


RESULT 62
US-09-949-016-198269
; Sequence 198269, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198269
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198269

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 295


RESULT 63
US-09-949-016-198270
; Sequence 198270, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 198270
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-198270

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db          60 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 105


RESULT 64
US-09-949-016-206260
; Sequence 206260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206260
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-206260

  Query Match             1.5%;  Score 46;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db          57 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 102


RESULT 65
US-09-949-016-206261
; Sequence 206261, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206261
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
```

US-09-949-016-206261

Query Match 1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||
Db 74 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 119

RESULT 66
US-09-949-016-206262
; Sequence 206262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206262
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-206262

Query Match 1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||
Db 90 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 135

RESULT 67
US-09-949-016-14878
; Sequence 14878, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14878
; LENGTH: 17704
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(17704)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14878

Query Match 1.5%; Score 46; DB 3; Length 17704;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||
Db 8184 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8229

RESULT 68
US-09-949-016-15290/c
; Sequence 15290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15290
; LENGTH: 27600
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(27600)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15290

Query Match 1.5%; Score 46; DB 3; Length 27600;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||
Db 3783 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 3738

RESULT 69
US-09-949-016-11978/c
; Sequence 11978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11978
; LENGTH: 29960
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11978

```
  Query Match             1.5%;  Score 46;  DB 3;  Length 29960;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       8547 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8502


RESULT 70
US-09-949-016-16100/c
; Sequence 16100, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16100
;   LENGTH: 30054
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16100

  Query Match             1.5%;  Score 46;  DB 3;  Length 30054;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       8640 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 8595


RESULT 71
US-09-949-016-16864/c
; Sequence 16864, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16864
;   LENGTH: 40905
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(40905)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16864

  Query Match             1.5%;  Score 46;  DB 3;  Length 40905;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      22092 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 22047


RESULT 72
US-09-949-016-12638/c
; Sequence 12638, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12638
;   LENGTH: 60917
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12638

  Query Match             1.5%;  Score 46;  DB 3;  Length 60917;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      39793 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 39748


RESULT 73
US-09-949-016-16057/c
; Sequence 16057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16057
;   LENGTH: 60917
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16057

  Query Match             1.5%;  Score 46;  DB 3;  Length 60917;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      39793 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 39748


RESULT 74
US-09-949-016-14736/c
; Sequence 14736, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14736
;   LENGTH: 84916
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(84916)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14736

  Query Match             1.5%;  Score 46;  DB 3;  Length 84916;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      15607 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 15562


RESULT 75
US-09-949-016-15606
; Sequence 15606, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15606
;   LENGTH: 96739
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(96739)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15606

  Query Match             1.5%;  Score 46;  DB 3;  Length 96739;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db      49090 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 49135


RESULT 76
US-09-949-016-17565/c
; Sequence 17565, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17565
;   LENGTH: 116955
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17565

  Query Match             1.5%;  Score 46;  DB 3;  Length 116955;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      52491 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 52446


RESULT 77
US-09-949-016-12439/c
; Sequence 12439, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12439
;   LENGTH: 118136
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(118136)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12439

  Query Match             1.5%;  Score 46;  DB 3;  Length 118136;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      15502 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 15457
```

RESULT 78
US-09-949-016-15316/c
; Sequence 15316, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15316
; LENGTH: 144322
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(144322)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15316

Query Match 1.5%; Score 46; DB 3; Length 144322;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      56738 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 56693
```

RESULT 79
US-09-949-016-16513/c
; Sequence 16513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16513
; LENGTH: 177293
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16513

Query Match 1.5%; Score 46; DB 3; Length 177293;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      34975 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 34930
```

RESULT 80
US-09-949-016-16264
; Sequence 16264, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16264
; LENGTH: 240157
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(240157)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16264

Query Match 1.5%; Score 46; DB 3; Length 240157;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db     184749 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 184794
```

RESULT 81
US-09-949-016-12656/c
; Sequence 12656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12656
; LENGTH: 253345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12656

Query Match 1.5%; Score 46; DB 3; Length 253345;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      62618 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 62573


RESULT 82
US-09-949-016-13639/c
; Sequence 13639, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13639
;   LENGTH: 253364
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13639

  Query Match             1.5%;  Score 46;  DB 3;  Length 253364;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      62618 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 62573


RESULT 83
US-09-949-016-12008/c
; Sequence 12008, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12008
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12008

  Query Match             1.5%;  Score 46;  DB 3;  Length 331814;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db     253007 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 252962


RESULT 84
US-09-949-016-17056/c
; Sequence 17056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17056
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17056

  Query Match             1.5%;  Score 46;  DB 3;  Length 331814;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db     253007 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 252962


RESULT 85
US-09-949-016-12147/c
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12147
;   LENGTH: 767677
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(767677)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

  Query Match             1.5%;  Score 46;  DB 3;  Length 767677;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-11;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db     681106 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 681061
```

RESULT 86
US-09-949-016-17361/c
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

Query Match 1.5%; Score 46; DB 3; Length 767677;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db     681106 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 681061
```

RESULT 87
US-09-621-976-15071
; Sequence 15071, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15071
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 250
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15071

Query Match 1.4%; Score 45; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        163 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 207
```

RESULT 88
US-09-949-016-30237
; Sequence 30237, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30237
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30237

Query Match 1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        535 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 579
```

RESULT 89
US-09-949-016-30238
; Sequence 30238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30238
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30238

Query Match 1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        401 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 445
```

RESULT 90
US-09-949-016-30239

```
; Sequence 30239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30239
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-30239

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        328 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 372


RESULT 91
US-09-949-016-65461/c
; Sequence 65461, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65461
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-65461

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        577 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 533


RESULT 92
US-09-949-016-72556
; Sequence 72556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72556
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-72556

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         71 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 115


RESULT 93
US-09-949-016-72557
; Sequence 72557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72557
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-72557

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        529 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 573


RESULT 94
US-09-949-016-82095
; Sequence 82095, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
```

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82095
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-82095

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        535 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 579


RESULT 95
US-09-949-016-82096
; Sequence 82096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82096
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-82096

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        401 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 445


RESULT 96
US-09-949-016-82097
; Sequence 82097, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82097
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-82097

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        328 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 372


RESULT 97
US-09-949-016-145830/c
; Sequence 145830, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 145830
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-145830

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        502 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 458


RESULT 98
US-09-949-016-146098/c
; Sequence 146098, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146098
;   LENGTH: 601
;   TYPE: DNA
```

```
;   ORGANISM: Human
US-09-949-016-146098

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        502 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 458


RESULT 99
US-09-949-016-146366/c
; Sequence 146366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146366
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-146366

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        502 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 458


RESULT 100
US-09-949-016-166442
; Sequence 166442, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 166442
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-166442

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||||
Db        377 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 421


RESULT 101
US-09-949-016-169130/c
; Sequence 169130, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169130
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169130

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||||
Db        519 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 475


RESULT 102
US-09-949-016-169131/c
; Sequence 169131, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169131
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169131

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
              |||||||||||||||||||||||||||||||||||||||||||||
Db        236 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 192
```

```
RESULT 103
US-09-949-016-203717
; Sequence 203717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203717
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-203717

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        535 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 579


RESULT 104
US-09-949-016-203718
; Sequence 203718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203718
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-203718

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        401 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 445


RESULT 105
US-09-949-016-203719
; Sequence 203719, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203719
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-203719

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        328 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 372


RESULT 106
US-09-949-002-4958
; Sequence 4958, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4958
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-4958

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        524 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 568


RESULT 107
US-09-949-002-4959
; Sequence 4959, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
```

```
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4959
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-4959

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        521 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 565


RESULT 108
US-09-949-002-9370/c
; Sequence 9370, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9370
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-9370

  Query Match             1.4%;  Score 45;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        489 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 445


RESULT 109
US-09-949-016-14483/c
; Sequence 14483, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14483
;   LENGTH: 8429
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14483

  Query Match             1.4%;  Score 45;  DB 3;  Length 8429;
  Best Local Similarity   100.0%;  Pred. No. 6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
              |||||||||||||||||||||||||||||||||||||||||||||
Db       4310 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 4266


RESULT 110
US-09-949-016-14852/c
; Sequence 14852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14852
;   LENGTH: 9123
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14852

  Query Match             1.4%;  Score 45;  DB 3;  Length 9123;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        664 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 620


RESULT 111
US-09-949-016-17432/c
; Sequence 17432, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17432
;   LENGTH: 9176
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17432

  Query Match             1.4%;  Score 45;  DB 3;  Length 9176;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-11;
```

```
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 620


RESULT 112
US-09-949-016-13766
; Sequence 13766, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13766
;   LENGTH: 10597
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13766

  Query Match             1.4%;  Score 45;  DB 3;  Length 10597;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        952 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 996


RESULT 113
US-09-949-016-17469/c
; Sequence 17469, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17469
;   LENGTH: 15975
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17469

  Query Match             1.4%;  Score 45;  DB 3;  Length 15975;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       8586 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 8542


RESULT 114
US-09-146-053-7
; Sequence 7, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
;  APPLICANT: Ryan, James W.
;  APPLICANT: Sprinkle, Terry Joe Curtis
;  APPLICANT: Venema, Richard C.
;  TITLE OF INVENTION: Human Aminopeptidase P Gene
;  FILE REFERENCE: MCG103
;  CURRENT APPLICATION NUMBER: US/09/146,053A
;  CURRENT FILING DATE: 1998-09-02
;  EARLIER APPLICATION NUMBER: 60/057,854
;  EARLIER FILING DATE: 1997-09-02
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 16595
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-146-053-7

  Query Match             1.4%;  Score 45;  DB 3;  Length 16595;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       4301 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4345


RESULT 115
US-09-949-016-16393/c
; Sequence 16393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16393
;   LENGTH: 19296
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16393

  Query Match             1.4%;  Score 45;  DB 3;  Length 19296;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db      14177 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 14133


RESULT 116
US-09-949-016-15920/c
; Sequence 15920, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
```

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15920
;   LENGTH: 19943
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15920

  Query Match             1.4%;  Score 45;  DB 3;  Length 19943;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db      18806 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 18762


RESULT 117
US-09-949-016-12553/c
; Sequence 12553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12553
;   LENGTH: 19945
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12553

  Query Match             1.4%;  Score 45;  DB 3;  Length 19945;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db      18806 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 18762


RESULT 118
US-09-949-016-13069/c
; Sequence 13069, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13069
;   LENGTH: 27545
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13069

  Query Match             1.4%;  Score 45;  DB 3;  Length 27545;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      25807 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 25763


RESULT 119
US-09-949-016-16094
; Sequence 16094, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16094
;   LENGTH: 33099
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(33099)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16094

  Query Match             1.4%;  Score 45;  DB 3;  Length 33099;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      32755 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 32799


RESULT 120
US-09-949-016-14417/c
; Sequence 14417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
```

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14417
;   LENGTH: 36223
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14417

  Query Match             1.4%;  Score 45;  DB 3;  Length 36223;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2978 GCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 3022
              |||||||||||||||||||||||||||||||||||||||||||||
Db       8552 GCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 8508


RESULT 121
US-09-949-016-17543/c
; Sequence 17543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17543
;   LENGTH: 36546
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17543

  Query Match             1.4%;  Score 45;  DB 3;  Length 36546;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db      27447 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 27403


RESULT 122
US-09-820-004-3
; Sequence 3, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201
;  CURRENT APPLICATION NUMBER: US/09/820,004
;  CURRENT FILING DATE:  2001-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 40090
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(40090)
;   OTHER INFORMATION: n = A,T,C or G
US-09-820-004-3

  Query Match             1.4%;  Score 45;  DB 3;  Length 40090;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              |||||||||||||||||||||||||||||||||||||||||||||
Db      25713 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 25757


RESULT 123
US-09-949-016-12374
; Sequence 12374, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12374
;   LENGTH: 40147
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12374

  Query Match             1.4%;  Score 45;  DB 3;  Length 40147;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              |||||||||||||||||||||||||||||||||||||||||||||
Db      25771 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 25815


RESULT 124
US-09-949-016-14128
; Sequence 14128, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14128
;   LENGTH: 40379
;   TYPE: DNA
;   ORGANISM: Human
```

US-09-949-016-14128

Query Match 1.4%; Score 45; DB 3; Length 40379;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||
Db 26306 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 26350

RESULT 125
US-09-949-016-17495
; Sequence 17495, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17495
; LENGTH: 40877
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17495

Query Match 1.4%; Score 45; DB 3; Length 40877;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||
Db 25815 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 25859

RESULT 126
US-09-949-002-845
; Sequence 845, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 845
; LENGTH: 40908
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-845

Query Match 1.4%; Score 45; DB 3; Length 40908;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
|||||||||||||||||||||||||||||||||||||||||||||
Db 16225 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 16269

RESULT 127
US-09-949-002-603
; Sequence 603, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 603
; LENGTH: 40944
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-603

Query Match 1.4%; Score 45; DB 3; Length 40944;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
|||||||||||||||||||||||||||||||||||||||||||||
Db 16262 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 16306

RESULT 128
US-09-949-016-16492
; Sequence 16492, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16492
; LENGTH: 47781
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47781)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16492

Query Match 1.4%; Score 45; DB 3; Length 47781;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
|||||||||||||||||||||||||||||||||||||||||||||
Db 46796 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 46840

RESULT 129

```
US-09-949-016-16493
; Sequence 16493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16493
;   LENGTH: 47781
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47781)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16493

  Query Match             1.4%;  Score 45;  DB 3;  Length 47781;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      46796 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 46840


RESULT 130
US-09-949-016-16494
; Sequence 16494, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16494
;   LENGTH: 47781
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47781)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16494

  Query Match             1.4%;  Score 45;  DB 3;  Length 47781;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      46796 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 46840


RESULT 131
US-09-949-002-817/c
; Sequence 817, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817
;   LENGTH: 49677
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-817

  Query Match             1.4%;  Score 45;  DB 3;  Length 49677;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      19485 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 19441


RESULT 132
US-09-949-016-15254
; Sequence 15254, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15254
;   LENGTH: 64984
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(64984)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15254

  Query Match             1.4%;  Score 45;  DB 3;  Length 64984;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      35279 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 35323


RESULT 133
US-09-949-016-12630/c
```

```
; Sequence 12630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12630
;   LENGTH: 66627
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12630

  Query Match             1.4%;  Score 45;  DB 3;  Length 66627;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      12900 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12856


RESULT 134
US-09-949-016-16112/c
; Sequence 16112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16112
;   LENGTH: 66628
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16112

  Query Match             1.4%;  Score 45;  DB 3;  Length 66628;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      12900 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12856


RESULT 135
US-09-949-002-714
; Sequence 714, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 714
;   LENGTH: 70313
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-714

  Query Match             1.4%;  Score 45;  DB 3;  Length 70313;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      37513 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 37557


RESULT 136
US-09-949-002-633
; Sequence 633, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633
;   LENGTH: 72347
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-633

  Query Match             1.4%;  Score 45;  DB 3;  Length 72347;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      37513 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 37557


RESULT 137
US-09-949-016-12659/c
; Sequence 12659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12659
;   LENGTH: 80858
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12659

  Query Match             1.4%;  Score 45;  DB 3;  Length 80858;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db      11289 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 11245


RESULT 138
US-09-949-016-15715/c
; Sequence 15715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15715
;   LENGTH: 80859
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15715

  Query Match             1.4%;  Score 45;  DB 3;  Length 80859;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db      11289 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 11245


RESULT 139
US-09-949-016-13655
; Sequence 13655, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13655
;   LENGTH: 87648
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(87648)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13655

  Query Match             1.4%;  Score 45;  DB 3;  Length 87648;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db      35917 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 35961


RESULT 140
US-09-949-016-15146
; Sequence 15146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15146
;   LENGTH: 91279
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15146

  Query Match             1.4%;  Score 45;  DB 3;  Length 91279;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db      68268 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 68312


RESULT 141
US-09-949-016-15944/c
; Sequence 15944, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15944
;   LENGTH: 93532
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15944
```

```
  Query Match             1.4%;  Score 45;  DB 3;  Length 93532;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      20699 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 20655


RESULT 142
US-09-949-016-16553
; Sequence 16553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16553
;   LENGTH: 94142
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(94142)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16553

  Query Match             1.4%;  Score 45;  DB 3;  Length 94142;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db      19351 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 19395


RESULT 143
US-09-949-016-17103
; Sequence 17103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17103
;   LENGTH: 96690
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17103

  Query Match             1.4%;  Score 45;  DB 3;  Length 96690;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db      72113 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 72157


RESULT 144
US-09-949-016-13319/c
; Sequence 13319, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13319
;   LENGTH: 103750
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13319

  Query Match             1.4%;  Score 45;  DB 3;  Length 103750;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||
Db      75631 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 75587


RESULT 145
US-09-949-016-16424
; Sequence 16424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16424
;   LENGTH: 112465
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16424

  Query Match             1.4%;  Score 45;  DB 3;  Length 112465;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||
Db      35617 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 35661


RESULT 146
US-09-949-016-13915
; Sequence 13915, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13915
;   LENGTH: 120609
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(120609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13915

  Query Match             1.4%;  Score 45;  DB 3;  Length 120609;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||
Db      83825 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 83869


RESULT 147
US-09-949-016-17074/c
; Sequence 17074, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17074
;   LENGTH: 140725
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17074

  Query Match             1.4%;  Score 45;  DB 3;  Length 140725;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||
Db      79776 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 79732


RESULT 148
US-09-949-016-13870/c
; Sequence 13870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13870
;   LENGTH: 155266
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13870

  Query Match             1.4%;  Score 45;  DB 3;  Length 155266;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||
Db      16826 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 16782


RESULT 149
US-09-949-016-13733
; Sequence 13733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13733
;   LENGTH: 162841
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13733

  Query Match             1.4%;  Score 45;  DB 3;  Length 162841;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
                |||||||||||||||||||||||||||||||||||||||||||||
Db     119589 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 119633


RESULT 150
US-09-949-016-16509/c
```

```
; Sequence 16509, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16509
;   LENGTH: 174639
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(174639)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16509

  Query Match             1.4%;  Score 45;  DB 3;  Length 174639;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 3115
             |||||||||||||||||||||||||||||||||||||||||||||
Db     107661 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT 107617


RESULT 151
US-09-949-016-15851
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15851
;   LENGTH: 205044
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(205044)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

  Query Match             1.4%;  Score 45;  DB 3;  Length 205044;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db     161139 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 161183


RESULT 152
US-09-949-016-15852
; Sequence 15852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15852
;   LENGTH: 205044
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(205044)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852

  Query Match             1.4%;  Score 45;  DB 3;  Length 205044;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db     161139 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 161183


RESULT 153
US-09-949-016-15853
; Sequence 15853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15853
;   LENGTH: 205044
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(205044)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15853

  Query Match             1.4%;  Score 45;  DB 3;  Length 205044;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||
Db     161139 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 161183


RESULT 154
US-09-949-016-17539/c
; Sequence 17539, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17539
;   LENGTH: 218940
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17539

  Query Match             1.4%;  Score 45;  DB 3;  Length 218940;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2978 GCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 3022
             |||||||||||||||||||||||||||||||||||||||||||||
Db     109636 GCCGGGCATGGTGGCGCATGCCTGTGGTCCCAGCTACTCGGGAGG 109592


RESULT 155
US-09-949-016-12387
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12387
;   LENGTH: 223471
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(223471)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387

  Query Match             1.4%;  Score 45;  DB 3;  Length 223471;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db     131566 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 131610


RESULT 156
US-09-949-016-12724
; Sequence 12724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12724
;   LENGTH: 223471
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(223471)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

  Query Match             1.4%;  Score 45;  DB 3;  Length 223471;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db     131566 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 131610


RESULT 157
US-09-949-016-12725
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12725
;   LENGTH: 223471
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(223471)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

  Query Match             1.4%;  Score 45;  DB 3;  Length 223471;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db     131566 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 131610


RESULT 158
US-09-513-999C-25910/c
; Sequence 25910, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 25910
;   LENGTH: 145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 120
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 121
;   OTHER INFORMATION: r=a or g
US-09-513-999C-25910

  Query Match             1.4%;  Score 44;  DB 3;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db         47 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 4


RESULT 159
US-09-513-999C-28305
; Sequence 28305, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 28305
;   LENGTH: 330
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-28305

  Query Match             1.4%;  Score 44;  DB 3;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              ||||||||||||||||||||||||||||||||||||||||||||
Db         65 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 108


RESULT 160
US-09-949-016-33113
; Sequence 33113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33113
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-33113

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db        430 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 473


RESULT 161
US-09-949-016-36581
; Sequence 36581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36581
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-36581

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 197
```

```
RESULT 162
US-09-949-016-38281
; Sequence 38281, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38281
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-38281

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        430 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 473


RESULT 163
US-09-949-016-46100
; Sequence 46100, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46100
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-46100

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 197


RESULT 164
US-09-949-016-46119
; Sequence 46119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46119
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-46119

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 197


RESULT 165
US-09-949-016-71064
; Sequence 71064, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71064
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71064

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db         59 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 102


RESULT 166
US-09-949-016-71065
; Sequence 71065, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
```

```
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71065
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71065

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        110 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 153


RESULT 167
US-09-949-016-71066
; Sequence 71066, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71066
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71066

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 369


RESULT 168
US-09-949-016-71067
; Sequence 71067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71067
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71067

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        359 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 402


RESULT 169
US-09-949-016-71068
; Sequence 71068, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71068
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71068

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        379 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 422


RESULT 170
US-09-949-016-71069
; Sequence 71069, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71069
```

```
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-71069

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        434 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 477


RESULT 171
US-09-949-016-121487/c
; Sequence 121487, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121487
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-121487

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        120 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 77


RESULT 172
US-09-949-016-121488/c
; Sequence 121488, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121488
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-121488

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 308


RESULT 173
US-09-949-016-133933/c
; Sequence 133933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133933
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-133933

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        370 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 327


RESULT 174
US-09-949-016-133934/c
; Sequence 133934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133934
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-133934

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
```

```
              ||||||||||||||||||||||||||||||||||||||||||||
Db        368 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 325


RESULT 175
US-09-949-016-143264
; Sequence 143264, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 143264
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-143264

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        233 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 276


RESULT 176
US-09-949-016-143435
; Sequence 143435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 143435
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-143435

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db        233 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 276


RESULT 177
US-09-949-016-144443/c
; Sequence 144443, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144443
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-144443

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db        158 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 115


RESULT 178
US-09-949-016-144444/c
; Sequence 144444, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144444
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-144444

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||
Db        412 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 369


RESULT 179
US-09-949-016-150959
; Sequence 150959, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
```

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150959
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-150959

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        233 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 276


RESULT 180
US-09-949-016-163060
; Sequence 163060, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163060
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-163060

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        158 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 201


RESULT 181
US-09-949-016-165548/c
; Sequence 165548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165548
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-165548

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        211 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 168


RESULT 182
US-09-949-016-199110/c
; Sequence 199110, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 199110
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-199110

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
Db        231 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 188


RESULT 183
US-09-949-016-202729
; Sequence 202729, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
```

```
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 202729
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-202729

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        231 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 274


RESULT 184
US-09-949-016-204472
; Sequence 204472, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204472
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-204472

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        430 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 473


RESULT 185
US-09-949-016-204542
; Sequence 204542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204542
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-204542

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        430 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 473


RESULT 186
US-09-949-002-10588/c
; Sequence 10588, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10588
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-002-10588

  Query Match             1.4%;  Score 44;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        168 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 125


RESULT 187
US-09-641-638-56
; Sequence 56, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 56
;    LENGTH: 956
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
;    FEATURE:
```

```
;   NAME/KEY: allele
;   LOCATION: 478
;   OTHER INFORMATION: 12-220-48 : polymorphic base G or A
;   NAME/KEY: misc_binding
;   LOCATION: 458..477
;   OTHER INFORMATION: 12-220-48.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 479..498
;   OTHER INFORMATION: 12-220-48.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 505..525
;   OTHER INFORMATION: upstream amplification primer, complement
;   NAME/KEY: primer_bind
;   LOCATION: 76..96
;   OTHER INFORMATION: downstream amplification primer
;   NAME/KEY: misc_binding
;   LOCATION: 466..490
;   OTHER INFORMATION: 12-220-48 potential probe
US-09-641-638-56

  Query Match             1.4%;  Score 44;  DB 3;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||
Db        802 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 845


RESULT 188
US-10-170-097-56
; Sequence 56, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE: GEN-T114XC2D1
;  CURRENT APPLICATION NUMBER: US/10/170,097
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/641,638
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: US 09/502,330
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: US 60/133,200
;  PRIOR FILING DATE: 1999-05-07
;  PRIOR APPLICATION NUMBER: US 09/275,267
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: US 60/119,917
;  PRIOR FILING DATE: 1999-02-12
;  NUMBER OF SEQ ID NOS: 1304
;  SOFTWARE: Patent.pm
; SEQ ID NO 56
;   LENGTH: 956
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 478
;   OTHER INFORMATION: 12-220-48 : polymorphic base G or A
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 458..477
;   OTHER INFORMATION: 12-220-48.mis1, potential
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 479..498
;   OTHER INFORMATION: 12-220-48.mis2, potential complement
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 505..525
;   OTHER INFORMATION: upstream amplification primer, complement
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 76..96
;   OTHER INFORMATION: downstream amplification primer
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 466..490
;   OTHER INFORMATION: 12-220-48 potential probe
US-10-170-097-56

  Query Match             1.4%;  Score 44;  DB 3;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||
Db        802 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 845


RESULT 189
US-09-949-016-11999
; Sequence 11999, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11999
;   LENGTH: 7480
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11999

  Query Match             1.4%;  Score 44;  DB 3;  Length 7480;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||
Db        247 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 290


RESULT 190
US-09-949-016-16486
; Sequence 16486, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16486
;   LENGTH: 7567
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16486

  Query Match             1.4%;  Score 44;  DB 3;  Length 7567;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        334 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 377


RESULT 191
US-09-318-448-5/c
; Sequence 5, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
;  APPLICANT: Johnson, William G.
;  APPLICANT: Stenroos, Edward S.
;  TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
;  TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
;  FILE REFERENCE: 601-1-057
;  CURRENT APPLICATION NUMBER: US/09/318,448
;  CURRENT FILING DATE: 1999-05-25
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 7720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-318-448-5

  Query Match             1.4%;  Score 44;  DB 3;  Length 7720;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db       5249 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 5206


RESULT 192
US-09-577-266-5/c
; Sequence 5, Application US/09577266
; Patent No. 6912492
; GENERAL INFORMATION:
;  APPLICANT:  Johnson, William G.
;  APPLICANT:  Stenroos, Edward S.
;  TITLE OF INVENTION:  METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
;  TITLE OF INVENTION:  DEVELOPMENTAL DISORDERS
;  FILE REFERENCE:  601-1-057N
;  CURRENT APPLICATION NUMBER: US/09/577,266
;  CURRENT FILING DATE:  2000-05-23
;  PRIOR APPLICATION NUMBER:  60/136,198
;  PRIOR FILING DATE:  1999-05-25
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 7720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-577-266-5

  Query Match             1.4%;  Score 44;  DB 3;  Length 7720;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db       5249 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 5206


RESULT 193
US-09-949-016-17443/c
; Sequence 17443, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17443
;   LENGTH: 10720
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17443

  Query Match             1.4%;  Score 44;  DB 3;  Length 10720;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db       3805 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3762


RESULT 194
US-09-949-016-12354
; Sequence 12354, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12354
;   LENGTH: 15116
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12354

  Query Match             1.4%;  Score 44;  DB 3;  Length 15116;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db       8745 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 8788


RESULT 195
US-09-949-016-16260
```

```
; Sequence 16260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16260
;   LENGTH: 15117
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16260

  Query Match             1.4%;  Score 44;  DB 3;  Length 15117;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2899 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              ||||||||||||||||||||||||||||||||||||||||||||
Db       8745 GATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 8788


RESULT 196
US-09-949-002-592
; Sequence 592, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 592
;   LENGTH: 18651
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-592

  Query Match             1.4%;  Score 44;  DB 3;  Length 18651;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db      14614 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14657


RESULT 197
US-09-949-002-786
; Sequence 786, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 786
;   LENGTH: 18682
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-786

  Query Match             1.4%;  Score 44;  DB 3;  Length 18682;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db      14634 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14677


RESULT 198
US-09-949-016-16523/c
; Sequence 16523, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16523
;   LENGTH: 18891
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(18891)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16523

  Query Match             1.4%;  Score 44;  DB 3;  Length 18891;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db      14166 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14123


RESULT 199
US-09-949-016-15113
; Sequence 15113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15113
; LENGTH: 19389
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15113

Query Match 1.4%; Score 44; DB 3; Length 19389;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db      14852 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14895
```

RESULT 200
US-09-949-016-15267
; Sequence 15267, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15267
; LENGTH: 19974
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(19974)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15267

Query Match 1.4%; Score 44; DB 3; Length 19974;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||
Db      15162 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 15205
```

Search completed: May 11, 2006, 06:00:44
Job time : 560 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:51:35 ; Search time 2382 Seconds
(without alignments)
10838.370 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 1

Total number of hits satisfying chosen parameters: 19584538

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	748	24.0	850	9	US-10-450-763-29532	Sequence 29532, A
c	2	657	21.0	708	6	US-10-029-386-20699	Sequence 20699, A
c	3	524	16.8	524	6	US-10-029-386-6988	Sequence 6988, Ap
c	4	509	16.3	560	4	US-09-925-065A-740956	Sequence 740956,
c	5	506	16.2	557	4	US-09-925-065A-769508	Sequence 769508,
c	6	505	16.2	556	4	US-09-925-065A-736351	Sequence 736351,
	7	504	16.1	555	4	US-09-925-065A-737120	Sequence 737120,
c	8	241	7.7	1349	9	US-10-450-763-8312	Sequence 8312, Ap
	9	60	1.9	60	3	US-09-908-975-23330	Sequence 23330, A
c	10	53	1.7	175737	8	US-10-723-860-1753	Sequence 1753, Ap
c	11	53	1.7	175737	9	US-10-783-271-34	Sequence 34, Appl
c	12	52	1.7	428	8	US-10-674-124A-76	Sequence 76, Appl
c	13	52	1.7	559	4	US-09-925-065A-591947	Sequence 591947,
	14	52	1.7	568	5	US-10-027-632-41688	Sequence 41688, A
	15	52	1.7	568	6	US-10-027-632-41688	Sequence 41688, A
	16	52	1.7	574	4	US-09-925-065A-768118	Sequence 768118,
c	17	52	1.7	592	4	US-09-925-065A-59682	Sequence 59682, A
c	18	52	1.7	617	4	US-09-925-065A-386361	Sequence 386361,
	19	52	1.7	631	5	US-10-027-632-61987	Sequence 61987, A
	20	52	1.7	631	5	US-10-027-632-61988	Sequence 61988, A
	21	52	1.7	631	5	US-10-027-632-61989	Sequence 61989, A
	22	52	1.7	631	6	US-10-027-632-61987	Sequence 61987, A
	23	52	1.7	631	6	US-10-027-632-61988	Sequence 61988, A
	24	52	1.7	631	6	US-10-027-632-61989	Sequence 61989, A
c	25	52	1.7	695	4	US-09-925-065A-708701	Sequence 708701,
c	26	52	1.7	1437	9	US-10-450-763-14141	Sequence 14141, A
c	27	52	1.7	9620	3	US-09-764-891-8895	Sequence 8895, Ap
	28	52	1.7	92139	3	US-09-918-686-1	Sequence 1, Appli
	29	52	1.7	92139	6	US-10-353-150-1	Sequence 1, Appli
c	30	52	1.7	130320	6	US-10-408-168-1	Sequence 1, Appli
	31	52	1.7	136726	6	US-10-085-117-244	Sequence 244, App
	32	52	1.7	150437	9	US-10-981-277-50	Sequence 50, Appl
	33	52	1.7	221000	6	US-10-174-014-12	Sequence 12, Appl
	34	52	1.7	233380	5	US-10-087-192-652	Sequence 652, App
c	35	52	1.7	398287	7	US-10-741-601-5719	Sequence 5719, Ap
c	36	52	1.7	398287	8	US-10-741-600-17839	Sequence 17839, A
c	37	51	1.6	385	4	US-09-925-065A-211249	Sequence 211249,
	38	51	1.6	561	4	US-09-925-065A-141833	Sequence 141833,
	39	51	1.6	622	5	US-10-027-632-54553	Sequence 54553, A
	40	51	1.6	622	5	US-10-027-632-55034	Sequence 55034, A
	41	51	1.6	622	5	US-10-027-632-56282	Sequence 56282, A
	42	51	1.6	622	5	US-10-027-632-56384	Sequence 56384, A
	43	51	1.6	622	5	US-10-027-632-180536	Sequence 180536,
	44	51	1.6	622	5	US-10-027-632-314485	Sequence 314485,
	45	51	1.6	622	6	US-10-027-632-54553	Sequence 54553, A
	46	51	1.6	622	6	US-10-027-632-55034	Sequence 55034, A
	47	51	1.6	622	6	US-10-027-632-56282	Sequence 56282, A
	48	51	1.6	622	6	US-10-027-632-56384	Sequence 56384, A
	49	51	1.6	622	6	US-10-027-632-180536	Sequence 180536,
	50	51	1.6	622	6	US-10-027-632-314485	Sequence 314485,
c	51	51	1.6	832	3	US-09-925-298-220	Sequence 220, App
c	52	51	1.6	832	5	US-10-102-806-220	Sequence 220, App
	53	51	1.6	2791	9	US-10-450-763-29534	Sequence 29534, A
c	54	51	1.6	23139	9	US-10-737-082-22	Sequence 22, Appl
c	55	51	1.6	23139	9	US-10-765-790-22	Sequence 22, Appl
c	56	51	1.6	68200	9	US-10-840-590-3	Sequence 3, Appli
c	57	51	1.6	215221	5	US-10-087-192-1360	Sequence 1360, Ap
	58	50	1.6	288	3	US-09-973-278-870	Sequence 870, App
	59	50	1.6	432	9	US-10-450-763-29529	Sequence 29529, A
c	60	50	1.6	571	4	US-09-925-065A-484883	Sequence 484883,
c	61	50	1.6	572	4	US-09-925-065A-785967	Sequence 785967,
c	62	50	1.6	625	4	US-09-925-065A-894777	Sequence 894777,
c	63	50	1.6	630	4	US-09-925-065A-892664	Sequence 892664,
c	64	50	1.6	633	4	US-09-925-065A-892405	Sequence 892405,
c	65	50	1.6	633	4	US-09-925-065A-892514	Sequence 892514,
	66	50	1.6	633	4	US-09-925-065A-949927	Sequence 949927,
	67	50	1.6	676	4	US-09-925-065A-917225	Sequence 917225,
	68	50	1.6	676	4	US-09-925-065A-917226	Sequence 917226,
	69	50	1.6	676	4	US-09-925-065A-944569	Sequence 944569,
c	70	50	1.6	714	5	US-10-027-632-158850	Sequence 158850,
c	71	50	1.6	714	6	US-10-027-632-158850	Sequence 158850,
	72	50	1.6	739	4	US-09-925-065A-934608	Sequence 934608,
	73	50	1.6	742	4	US-09-925-065A-941367	Sequence 941367,
	74	50	1.6	1986	5	US-10-027-632-265269	Sequence 265269,
	75	50	1.6	1986	6	US-10-027-632-265269	Sequence 265269,
	76	50	1.6	8705	6	US-10-374-979-10	Sequence 10, Appl
	77	50	1.6	8705	7	US-10-182-936A-10	Sequence 10, Appl
	78	50	1.6	8705	8	US-10-731-739-10	Sequence 10, Appl
	79	50	1.6	8705	8	US-10-477-238A-10	Sequence 10, Appl
	80	50	1.6	8705	8	US-10-680-287A-10	Sequence 10, Appl
	81	50	1.6	8705	9	US-10-477-173-10	Sequence 10, Appl
	82	50	1.6	8705	9	US-10-834-377-10	Sequence 10, Appl
c	83	50	1.6	31474	3	US-09-764-891-8149	Sequence 8149, Ap
c	84	50	1.6	32189	3	US-09-764-878-379	Sequence 379, App
c	85	50	1.6	32189	5	US-10-079-854-379	Sequence 379, App
c	86	50	1.6	32193	3	US-09-764-877-2623	Sequence 2623, Ap
c	87	50	1.6	32193	6	US-10-216-464-37	Sequence 37, Appl
c	88	50	1.6	32193	6	US-10-242-515-2623	Sequence 2623, Ap
c	89	50	1.6	32221	3	US-09-764-878-377	Sequence 377, App
c	90	50	1.6	32221	5	US-10-079-854-377	Sequence 377, App
	91	50	1.6	33126	6	US-10-085-117-190	Sequence 190, App
	92	50	1.6	56510	8	US-10-741-600-17637	Sequence 17637, A
c	93	50	1.6	66973	5	US-10-087-192-574	Sequence 574, App
	94	50	1.6	156843	5	US-10-087-192-1408	Sequence 1408, Ap
	95	50	1.6	202814	8	US-10-719-993-6812	Sequence 6812, Ap
	96	50	1.6	276820	7	US-10-271-416-9	Sequence 9, Appli

c 97	49	1.6	420	4	US-09-925-065A-140670	Sequence 140670,
c 98	49	1.6	532	4	US-09-925-065A-140671	Sequence 140671,
c 99	49	1.6	552	4	US-09-925-065A-18609	Sequence 18609, A
c 100	49	1.6	552	4	US-09-925-065A-18610	Sequence 18610, A
c 101	49	1.6	552	4	US-09-925-065A-18611	Sequence 18611, A
c 102	49	1.6	580	5	US-10-027-632-80386	Sequence 80386, A
c 103	49	1.6	580	5	US-10-027-632-80387	Sequence 80387, A
c 104	49	1.6	580	6	US-10-027-632-80386	Sequence 80386, A
c 105	49	1.6	580	6	US-10-027-632-80387	Sequence 80387, A
c 106	49	1.6	695	4	US-09-925-065A-708700	Sequence 708700,
c 107	49	1.6	750	4	US-09-925-065A-932035	Sequence 932035,
c 108	49	1.6	750	4	US-09-925-065A-932036	Sequence 932036,
109	49	1.6	3170	5	US-10-027-632-113573	Sequence 113573,
110	49	1.6	3170	5	US-10-027-632-113574	Sequence 113574,
111	49	1.6	3170	6	US-10-027-632-113573	Sequence 113573,
112	49	1.6	3170	6	US-10-027-632-113574	Sequence 113574,
113	48	1.5	460	5	US-10-027-632-182264	Sequence 182264,
114	48	1.5	460	5	US-10-027-632-319423	Sequence 319423,
115	48	1.5	460	6	US-10-027-632-182264	Sequence 182264,
116	48	1.5	460	6	US-10-027-632-319423	Sequence 319423,
117	48	1.5	716	5	US-10-027-632-132626	Sequence 132626,
118	48	1.5	716	6	US-10-027-632-132626	Sequence 132626,
119	48	1.5	52242	7	US-10-052-482-172	Sequence 172, App
c 120	48	1.5	53779	9	US-10-737-082-85	Sequence 85, Appl
c 121	48	1.5	53779	9	US-10-765-790-85	Sequence 85, Appl
122	48	1.5	141121	8	US-10-741-600-17651	Sequence 17651, A
123	48	1.5	181684	5	US-10-087-192-790	Sequence 790, App
124	47	1.5	440	8	US-10-357-930-16322	Sequence 16322, A
125	47	1.5	516	8	US-10-357-930-46148	Sequence 46148, A
c 126	47	1.5	534	4	US-09-925-065A-440136	Sequence 440136,
c 127	47	1.5	534	4	US-09-925-065A-440137	Sequence 440137,
c 128	47	1.5	566	4	US-09-925-065A-623360	Sequence 623360,
c 129	47	1.5	3231	5	US-10-027-632-115306	Sequence 115306,
c 130	47	1.5	3231	5	US-10-027-632-115307	Sequence 115307,
c 131	47	1.5	3231	5	US-10-027-632-115308	Sequence 115308,
c 132	47	1.5	3231	6	US-10-027-632-115306	Sequence 115306,
c 133	47	1.5	3231	6	US-10-027-632-115307	Sequence 115307,
c 134	47	1.5	3231	6	US-10-027-632-115308	Sequence 115308,
c 135	46	1.5	201	8	US-10-719-993-51173	Sequence 51173, A
c 136	46	1.5	322	8	US-10-674-124A-25317	Sequence 25317, A
c 137	46	1.5	380	8	US-10-914-037-282	Sequence 282, App
c 138	46	1.5	405	3	US-09-867-701-6376	Sequence 6376, Ap
c 139	46	1.5	440	8	US-10-674-124A-24232	Sequence 24232, A
c 140	46	1.5	458	3	US-09-814-353-17260	Sequence 17260, A
141	46	1.5	482	8	US-10-674-124A-16717	Sequence 16717, A
142	46	1.5	497	8	US-10-357-930-60554	Sequence 60554, A
143	46	1.5	535	5	US-10-027-632-91304	Sequence 91304, A
144	46	1.5	535	5	US-10-027-632-91305	Sequence 91305, A
145	46	1.5	535	5	US-10-027-632-91306	Sequence 91306, A
146	46	1.5	535	6	US-10-027-632-91304	Sequence 91304, A
147	46	1.5	535	6	US-10-027-632-91305	Sequence 91305, A
148	46	1.5	535	6	US-10-027-632-91306	Sequence 91306, A
149	46	1.5	559	4	US-09-925-065A-124249	Sequence 124249,
c 150	46	1.5	606	4	US-09-925-065A-843811	Sequence 843811,
c 151	46	1.5	609	4	US-09-925-065A-801712	Sequence 801712,
c 152	46	1.5	610	4	US-09-925-065A-48663	Sequence 48663, A
153	46	1.5	610	4	US-09-925-065A-65044	Sequence 65044, A
c 154	46	1.5	614	4	US-09-925-065A-47410	Sequence 47410, A
c 155	46	1.5	614	4	US-09-925-065A-47411	Sequence 47411, A
c 156	46	1.5	615	4	US-09-925-065A-930543	Sequence 930543,
c 157	46	1.5	615	4	US-09-925-065A-930544	Sequence 930544,
c 158	46	1.5	616	4	US-09-925-065A-870312	Sequence 870312,
159	46	1.5	621	5	US-10-027-632-115128	Sequence 115128,
160	46	1.5	621	5	US-10-027-632-115129	Sequence 115129,
161	46	1.5	621	6	US-10-027-632-115128	Sequence 115128,
162	46	1.5	621	6	US-10-027-632-115129	Sequence 115129,
163	46	1.5	654	4	US-09-925-065A-554324	Sequence 554324,
c 164	46	1.5	654	4	US-09-925-065A-759345	Sequence 759345,
165	46	1.5	659	4	US-09-925-065A-790726	Sequence 790726,
166	46	1.5	659	4	US-09-925-065A-790727	Sequence 790727,
167	46	1.5	661	4	US-09-925-065A-771454	Sequence 771454,
168	46	1.5	672	4	US-09-925-065A-737520	Sequence 737520,
c 169	46	1.5	680	4	US-09-925-065A-737227	Sequence 737227,
170	46	1.5	689	4	US-09-925-065A-852555	Sequence 852555,
171	46	1.5	689	4	US-09-925-065A-852556	Sequence 852556,
c 172	46	1.5	756	5	US-10-027-632-144717	Sequence 144717,
c 173	46	1.5	756	5	US-10-027-632-144718	Sequence 144718,
c 174	46	1.5	756	5	US-10-027-632-144719	Sequence 144719,
c 175	46	1.5	756	6	US-10-027-632-144717	Sequence 144717,
c 176	46	1.5	756	6	US-10-027-632-144718	Sequence 144718,
c 177	46	1.5	756	6	US-10-027-632-144719	Sequence 144719,
c 178	46	1.5	840	4	US-09-925-065A-938958	Sequence 938958,
c 179	46	1.5	840	4	US-09-925-065A-938959	Sequence 938959,
c 180	46	1.5	840	4	US-09-925-065A-938960	Sequence 938960,
181	46	1.5	1105	4	US-09-925-065A-711495	Sequence 711495,
182	46	1.5	1364	4	US-09-925-065A-678054	Sequence 678054,
183	46	1.5	1635	4	US-09-925-065A-689300	Sequence 689300,
184	46	1.5	2227	4	US-09-925-065A-677066	Sequence 677066,
185	46	1.5	2798	5	US-10-027-632-260094	Sequence 260094,
186	46	1.5	2798	6	US-10-027-632-260094	Sequence 260094,
c 187	46	1.5	7001	6	US-10-008-789-10	Sequence 10, Appl
c 188	46	1.5	7001	10	US-11-071-724-10	Sequence 10, Appl
189	46	1.5	13409	3	US-09-764-891-9601	Sequence 9601, Ap
190	46	1.5	13409	5	US-10-205-428-930	Sequence 930, App
191	46	1.5	18501	3	US-09-764-847-1916	Sequence 1916, Ap
192	46	1.5	18501	5	US-10-092-154-1916	Sequence 1916, Ap
c 193	46	1.5	25001	6	US-10-187-659A-11	Sequence 11, Appl
c 194	46	1.5	28616	7	US-10-455-552-1	Sequence 1, Appli
c 195	46	1.5	57095	8	US-10-719-993-7030	Sequence 7030, Ap
c 196	46	1.5	58922	7	US-10-322-281-526	Sequence 526, App
c 197	46	1.5	59914	7	US-10-741-601-5619	Sequence 5619, Ap
c 198	46	1.5	93011	8	US-10-719-993-6871	Sequence 6871, Ap
199	46	1.5	96595	7	US-10-052-482-232	Sequence 232, App
c 200	46	1.5	99100	9	US-10-704-513-1	Sequence 1, Appli
c 201	46	1.5	99250	9	US-10-840-590-4	Sequence 4, Appli
202	46	1.5	101782	7	US-10-741-601-5661	Sequence 5661, Ap
c 203	46	1.5	103391	9	US-10-981-277-54	Sequence 54, Appl
204	46	1.5	141463	5	US-10-087-192-22	Sequence 22, Appl
205	46	1.5	168821	5	US-10-087-192-622	Sequence 622, App
206	46	1.5	181259	7	US-10-456-930-2	Sequence 2, Appli
c 207	46	1.5	193357	6	US-10-085-117-142	Sequence 142, App
208	46	1.5	256157	5	US-10-087-192-1204	Sequence 1204, Ap
209	46	1.5	256157	7	US-10-322-281-776	Sequence 776, App
210	46	1.5	276276	5	US-10-087-192-754	Sequence 754, App
c 211	46	1.5	347814	7	US-10-322-696-76	Sequence 76, Appl
c 212	46	1.5	410846	9	US-10-481-613-1	Sequence 1, Appli
c 213	46	1.5	717651	8	US-10-719-993-6817	Sequence 6817, Ap
c 214	46	1.5	744802	6	US-10-292-798-1369	Sequence 1369, Ap
c 215	45	1.4	145	3	US-09-764-868-1298	Sequence 1298, Ap
216	45	1.4	145	3	US-09-764-875-1230	Sequence 1230, Ap
217	45	1.4	425	5	US-10-027-632-290871	Sequence 290871,
218	45	1.4	425	6	US-10-027-632-290871	Sequence 290871,
219	45	1.4	492	3	US-09-918-995-24329	Sequence 24329, A
c 220	45	1.4	541	5	US-10-027-632-193287	Sequence 193287,
c 221	45	1.4	541	5	US-10-027-632-193288	Sequence 193288,
c 222	45	1.4	541	6	US-10-027-632-193287	Sequence 193287,
c 223	45	1.4	541	6	US-10-027-632-193288	Sequence 193288,
224	45	1.4	545	4	US-09-925-065A-765896	Sequence 765896,
c 225	45	1.4	546	4	US-09-925-065A-799095	Sequence 799095,
226	45	1.4	549	4	US-09-925-065A-615308	Sequence 615308,
c 227	45	1.4	552	5	US-10-027-632-68940	Sequence 68940, A
c 228	45	1.4	552	5	US-10-027-632-294750	Sequence 294750,
c 229	45	1.4	552	6	US-10-027-632-68940	Sequence 68940, A
c 230	45	1.4	552	6	US-10-027-632-294750	Sequence 294750,
231	45	1.4	558	5	US-10-027-632-288369	Sequence 288369,
232	45	1.4	558	6	US-10-027-632-288369	Sequence 288369,
c 233	45	1.4	559	4	US-09-925-065A-591948	Sequence 591948,
c 234	45	1.4	563	4	US-09-925-065A-12438	Sequence 12438, A
235	45	1.4	568	4	US-09-925-065A-13098	Sequence 13098, A
236	45	1.4	568	4	US-09-925-065A-338452	Sequence 338452,
c 237	45	1.4	590	4	US-09-925-065A-371064	Sequence 371064,
238	45	1.4	592	4	US-09-925-065A-927764	Sequence 927764,
239	45	1.4	592	4	US-09-925-065A-927765	Sequence 927765,
240	45	1.4	592	4	US-09-925-065A-949474	Sequence 949474,
c 241	45	1.4	596	4	US-09-925-065A-695847	Sequence 695847,
c 242	45	1.4	598	4	US-09-925-065A-10892	Sequence 10892, A

	243	45	1.4	601	9	US-10-893-315-1202	Sequence 1202, Ap
c	244	45	1.4	604	4	US-09-925-065A-160227	Sequence 160227,
c	245	45	1.4	604	4	US-09-925-065A-799096	Sequence 799096,
c	246	45	1.4	605	4	US-09-925-065A-795229	Sequence 795229,
c	247	45	1.4	605	4	US-09-925-065A-850785	Sequence 850785,
c	248	45	1.4	612	5	US-10-027-632-295528	Sequence 295528,
c	249	45	1.4	612	6	US-10-027-632-295528	Sequence 295528,
c	250	45	1.4	624	4	US-09-925-065A-846491	Sequence 846491,
c	251	45	1.4	627	4	US-09-925-065A-839499	Sequence 839499,
	252	45	1.4	629	5	US-10-027-632-14053	Sequence 14053, A
	253	45	1.4	629	6	US-10-027-632-14053	Sequence 14053, A
c	254	45	1.4	642	5	US-10-027-632-105085	Sequence 105085,
c	255	45	1.4	642	5	US-10-027-632-135498	Sequence 135498,
c	256	45	1.4	642	5	US-10-027-632-135499	Sequence 135499,
c	257	45	1.4	642	5	US-10-027-632-135500	Sequence 135500,
c	258	45	1.4	642	6	US-10-027-632-105085	Sequence 105085,
c	259	45	1.4	642	6	US-10-027-632-135498	Sequence 135498,
c	260	45	1.4	642	6	US-10-027-632-135499	Sequence 135499,
c	261	45	1.4	642	6	US-10-027-632-135500	Sequence 135500,
	262	45	1.4	652	5	US-10-027-632-184845	Sequence 184845,
	263	45	1.4	652	5	US-10-027-632-184846	Sequence 184846,
	264	45	1.4	652	6	US-10-027-632-184845	Sequence 184845,
	265	45	1.4	652	6	US-10-027-632-184846	Sequence 184846,
	266	45	1.4	660	5	US-10-027-632-216907	Sequence 216907,
	267	45	1.4	660	5	US-10-027-632-216908	Sequence 216908,
	268	45	1.4	660	6	US-10-027-632-216907	Sequence 216907,
	269	45	1.4	660	6	US-10-027-632-216908	Sequence 216908,
c	270	45	1.4	662	4	US-09-925-065A-675355	Sequence 675355,
c	271	45	1.4	662	4	US-09-925-065A-675356	Sequence 675356,
	272	45	1.4	677	5	US-10-027-632-121998	Sequence 121998,
	273	45	1.4	677	6	US-10-027-632-121998	Sequence 121998,
	274	45	1.4	702	5	US-10-027-632-145794	Sequence 145794,
	275	45	1.4	702	6	US-10-027-632-145794	Sequence 145794,
c	276	45	1.4	705	5	US-10-027-632-149399	Sequence 149399,
c	277	45	1.4	705	6	US-10-027-632-149399	Sequence 149399,
c	278	45	1.4	793	4	US-09-925-065A-13524	Sequence 13524, A
c	279	45	1.4	793	4	US-09-925-065A-77588	Sequence 77588, A
c	280	45	1.4	794	5	US-10-027-632-36618	Sequence 36618, A
c	281	45	1.4	794	6	US-10-027-632-36618	Sequence 36618, A
	282	45	1.4	896	4	US-09-925-065A-88887	Sequence 88887, A
	283	45	1.4	1027	4	US-09-925-065A-85535	Sequence 85535, A
	284	45	1.4	1027	4	US-09-925-065A-85536	Sequence 85536, A
c	285	45	1.4	1302	4	US-09-925-065A-553464	Sequence 553464,
	286	45	1.4	1305	4	US-09-925-065A-34751	Sequence 34751, A
	287	45	1.4	1498	4	US-09-925-065A-71613	Sequence 71613, A
c	288	45	1.4	1792	4	US-09-925-065A-549910	Sequence 549910,
c	289	45	1.4	1869	4	US-09-925-065A-689963	Sequence 689963,
c	290	45	1.4	4513	6	US-10-264-237-2853	Sequence 2853, Ap
c	291	45	1.4	6565	3	US-09-764-891-5729	Sequence 5729, Ap
c	292	45	1.4	6565	3	US-09-764-891-5730	Sequence 5730, Ap
	293	45	1.4	17581	3	US-09-764-869-2170	Sequence 2170, Ap
	294	45	1.4	17581	5	US-10-091-504-2170	Sequence 2170, Ap
	295	45	1.4	17581	6	US-10-227-577-2170	Sequence 2170, Ap
	296	45	1.4	17946	3	US-09-764-869-2312	Sequence 2312, Ap
	297	45	1.4	17946	5	US-10-091-504-2312	Sequence 2312, Ap
	298	45	1.4	17946	6	US-10-227-577-2312	Sequence 2312, Ap
	299	45	1.4	23694	5	US-10-087-192-1936	Sequence 1936, Ap
c	300	45	1.4	28693	8	US-10-741-600-17761	Sequence 17761, A
	301	45	1.4	31116	5	US-10-087-192-1660	Sequence 1660, Ap
	302	45	1.4	32706	5	US-10-087-192-1750	Sequence 1750, Ap
c	303	45	1.4	36534	6	US-10-285-351B-3	Sequence 3, Appli
c	304	45	1.4	36534	7	US-10-240-425-1461	Sequence 1461, Ap
c	305	45	1.4	38678	9	US-10-893-315-136	Sequence 136, App
c	306	45	1.4	38684	9	US-10-893-315-154	Sequence 154, App
	307	45	1.4	39566	7	US-10-304-125-11	Sequence 11, Appl
	308	45	1.4	40090	3	US-09-820-004-3	Sequence 3, Appli
	309	45	1.4	40090	7	US-10-644-021A-3	Sequence 3, Appli
	310	45	1.4	65277	9	US-10-461-862-152	Sequence 152, App
c	311	45	1.4	73995	5	US-10-087-192-208	Sequence 208, App
c	312	45	1.4	89900	9	US-10-722-939-3	Sequence 3, Appli
	313	45	1.4	103574	9	US-10-756-149-2307	Sequence 2307, Ap
c	314	45	1.4	107543	7	US-10-322-281-706	Sequence 706, App
c	315	45	1.4	107745	7	US-10-322-281-268	Sequence 268, App
	316	45	1.4	115755	8	US-10-719-993-6946	Sequence 6946, Ap
	317	45	1.4	116561	8	US-10-723-860-409	Sequence 409, App
c	318	45	1.4	149671	6	US-10-236-031B-53	Sequence 53, Appl
c	319	45	1.4	150573	9	US-10-981-277-56	Sequence 56, Appl
	320	45	1.4	191584	7	US-10-322-281-800	Sequence 800, App
	321	45	1.4	300000	6	US-10-262-552-33	Sequence 33, Appl
	322	45	1.4	300000	7	US-10-703-210-33	Sequence 33, Appl
c	323	45	1.4	366803	8	US-10-719-993-6805	Sequence 6805, Ap
	324	45	1.4	368004	3	US-09-949-654-3	Sequence 3, Appli
c	325	45	1.4	567564	8	US-10-699-156-3	Sequence 3, Appli
c	326	45	1.4	786452	8	US-10-719-993-6822	Sequence 6822, Ap
c	327	45	1.4	1691139	5	US-10-067-514-1	Sequence 1, Appli
c	328	45	1.4	1691139	7	US-10-419-723-1	Sequence 1, Appli
c	329	45	1.4	1691139	7	US-10-255-120-1	Sequence 1, Appli
c	330	45	1.4	1691140	9	US-10-868-397-1	Sequence 1, Appli
	331	44	1.4	98	9	US-10-708-204-6436	Sequence 6436, Ap
	332	44	1.4	116	3	US-09-984-429-588	Sequence 588, App
	333	44	1.4	129	7	US-10-242-535A-52981	Sequence 52981, A
	334	44	1.4	129	7	US-10-085-783A-52981	Sequence 52981, A
	335	44	1.4	182	3	US-09-764-887-597	Sequence 597, App
	336	44	1.4	182	5	US-10-073-961-597	Sequence 597, App
	337	44	1.4	201	8	US-10-719-993-12850	Sequence 12850, A
	338	44	1.4	250	5	US-10-027-632-22856	Sequence 22856, A
	339	44	1.4	250	6	US-10-027-632-22856	Sequence 22856, A
c	340	44	1.4	293	8	US-10-674-124A-23859	Sequence 23859, A
	341	44	1.4	312	8	US-10-674-124A-932	Sequence 932, App
c	342	44	1.4	423	4	US-09-925-065A-566681	Sequence 566681,
c	343	44	1.4	423	4	US-09-925-065A-566682	Sequence 566682,
c	344	44	1.4	423	5	US-10-027-632-36530	Sequence 36530, A
c	345	44	1.4	423	5	US-10-027-632-36531	Sequence 36531, A
	346	44	1.4	423	5	US-10-027-632-80045	Sequence 80045, A
	347	44	1.4	423	5	US-10-027-632-80046	Sequence 80046, A
	348	44	1.4	423	5	US-10-027-632-109591	Sequence 109591,
	349	44	1.4	423	5	US-10-027-632-109592	Sequence 109592,
c	350	44	1.4	423	6	US-10-027-632-36530	Sequence 36530, A
c	351	44	1.4	423	6	US-10-027-632-36531	Sequence 36531, A
	352	44	1.4	423	6	US-10-027-632-80045	Sequence 80045, A
	353	44	1.4	423	6	US-10-027-632-80046	Sequence 80046, A
	354	44	1.4	423	6	US-10-027-632-109591	Sequence 109591,
	355	44	1.4	423	6	US-10-027-632-109592	Sequence 109592,
	356	44	1.4	426	3	US-09-954-456-2027	Sequence 2027, Ap
	357	44	1.4	426	9	US-10-843-641A-5054	Sequence 5054, Ap
c	358	44	1.4	464	4	US-09-925-065A-602339	Sequence 602339,
	359	44	1.4	470	4	US-09-925-065A-767971	Sequence 767971,
c	360	44	1.4	470	8	US-10-674-124A-9871	Sequence 9871, Ap
c	361	44	1.4	470	8	US-10-674-124A-20267	Sequence 20267, A
c	362	44	1.4	496	4	US-09-925-065A-876040	Sequence 876040,
c	363	44	1.4	496	4	US-09-925-065A-876041	Sequence 876041,
c	364	44	1.4	496	4	US-09-925-065A-876042	Sequence 876042,
c	365	44	1.4	496	4	US-09-925-065A-905372	Sequence 905372,
	366	44	1.4	530	4	US-09-925-065A-297221	Sequence 297221,
	367	44	1.4	530	4	US-09-925-065A-297222	Sequence 297222,
	368	44	1.4	537	4	US-09-925-065A-448591	Sequence 448591,
	369	44	1.4	537	4	US-09-925-065A-448592	Sequence 448592,
c	370	44	1.4	556	4	US-09-925-065A-109914	Sequence 109914,
c	371	44	1.4	557	4	US-09-925-065A-870384	Sequence 870384,
c	372	44	1.4	557	4	US-09-925-065A-870385	Sequence 870385,
	373	44	1.4	562	4	US-09-925-065A-78244	Sequence 78244, A
c	374	44	1.4	565	5	US-10-027-632-127998	Sequence 127998,
c	375	44	1.4	565	6	US-10-027-632-127998	Sequence 127998,
	376	44	1.4	577	4	US-09-925-065A-707681	Sequence 707681,
	377	44	1.4	577	4	US-09-925-065A-707682	Sequence 707682,
c	378	44	1.4	585	5	US-10-027-632-201634	Sequence 201634,
c	379	44	1.4	585	6	US-10-027-632-201634	Sequence 201634,
	380	44	1.4	588	4	US-09-925-065A-729173	Sequence 729173,
c	381	44	1.4	589	4	US-09-925-065A-914465	Sequence 914465,
c	382	44	1.4	589	5	US-10-027-632-197271	Sequence 197271,
c	383	44	1.4	589	5	US-10-027-632-197272	Sequence 197272,
c	384	44	1.4	589	6	US-10-027-632-197271	Sequence 197271,
c	385	44	1.4	589	6	US-10-027-632-197272	Sequence 197272,
c	386	44	1.4	590	5	US-10-027-632-94737	Sequence 94737, A
c	387	44	1.4	590	5	US-10-027-632-318528	Sequence 318528,
c	388	44	1.4	590	6	US-10-027-632-94737	Sequence 94737, A

c 389	44	1.4	590	6	US-10-027-632-318528	Sequence 318528,
c 390	44	1.4	591	4	US-09-925-065A-936386	Sequence 936386,
c 391	44	1.4	593	4	US-09-925-065A-237003	Sequence 237003,
c 392	44	1.4	593	4	US-09-925-065A-237005	Sequence 237005,
c 393	44	1.4	595	4	US-09-925-065A-685243	Sequence 685243,
c 394	44	1.4	599	4	US-09-925-065A-902996	Sequence 902996,
395	44	1.4	600	9	US-10-956-157-7469	Sequence 7469, Ap
396	44	1.4	611	7	US-10-276-774-491	Sequence 491, App
397	44	1.4	612	4	US-09-925-065A-21972	Sequence 21972, A
c 398	44	1.4	614	5	US-10-027-632-260590	Sequence 260590,
c 399	44	1.4	614	6	US-10-027-632-260590	Sequence 260590,
c 400	44	1.4	618	4	US-09-925-065A-697941	Sequence 697941,
c 401	44	1.4	618	4	US-09-925-065A-697942	Sequence 697942,
c 402	44	1.4	618	4	US-09-925-065A-697943	Sequence 697943,
c 403	44	1.4	619	4	US-09-925-065A-789262	Sequence 789262,
c 404	44	1.4	619	4	US-09-925-065A-789263	Sequence 789263,
c 405	44	1.4	619	4	US-09-925-065A-789264	Sequence 789264,
c 406	44	1.4	619	4	US-09-925-065A-847245	Sequence 847245,
c 407	44	1.4	619	4	US-09-925-065A-951338	Sequence 951338,
408	44	1.4	623	4	US-09-925-065A-521713	Sequence 521713,
c 409	44	1.4	627	4	US-09-925-065A-545865	Sequence 545865,
c 410	44	1.4	627	4	US-09-925-065A-545866	Sequence 545866,
c 411	44	1.4	627	4	US-09-925-065A-545867	Sequence 545867,
c 412	44	1.4	627	4	US-09-925-065A-923432	Sequence 923432,
413	44	1.4	629	4	US-09-925-065A-821309	Sequence 821309,
c 414	44	1.4	637	5	US-10-027-632-187929	Sequence 187929,
c 415	44	1.4	637	6	US-10-027-632-187929	Sequence 187929,
416	44	1.4	643	4	US-09-925-065A-692017	Sequence 692017,
c 417	44	1.4	659	4	US-09-925-065A-741517	Sequence 741517,
c 418	44	1.4	659	4	US-09-925-065A-741518	Sequence 741518,
419	44	1.4	668	4	US-09-925-065A-127024	Sequence 127024,
420	44	1.4	673	4	US-09-925-065A-32538	Sequence 32538, A
c 421	44	1.4	698	5	US-10-027-632-252317	Sequence 252317,
c 422	44	1.4	698	6	US-10-027-632-252317	Sequence 252317,
423	44	1.4	738	5	US-10-027-632-139602	Sequence 139602,
c 424	44	1.4	738	5	US-10-027-632-172816	Sequence 172816,
425	44	1.4	738	6	US-10-027-632-139602	Sequence 139602,
c 426	44	1.4	738	6	US-10-027-632-172816	Sequence 172816,
c 427	44	1.4	748	5	US-10-027-632-146061	Sequence 146061,
c 428	44	1.4	748	6	US-10-027-632-146061	Sequence 146061,
c 429	44	1.4	751	5	US-10-027-632-14193	Sequence 14193, A
c 430	44	1.4	751	6	US-10-027-632-14193	Sequence 14193, A
c 431	44	1.4	755	5	US-10-027-632-27592	Sequence 27592, A
c 432	44	1.4	755	5	US-10-027-632-27593	Sequence 27593, A
c 433	44	1.4	755	6	US-10-027-632-27592	Sequence 27592, A
c 434	44	1.4	755	6	US-10-027-632-27593	Sequence 27593, A
c 435	44	1.4	756	5	US-10-027-632-24620	Sequence 24620, A
c 436	44	1.4	756	5	US-10-027-632-149995	Sequence 149995,
c 437	44	1.4	756	6	US-10-027-632-24620	Sequence 24620, A
c 438	44	1.4	756	6	US-10-027-632-149995	Sequence 149995,
c 439	44	1.4	784	4	US-09-925-065A-927249	Sequence 927249,
440	44	1.4	810	5	US-10-027-632-154967	Sequence 154967,
441	44	1.4	810	6	US-10-027-632-154967	Sequence 154967,
c 442	44	1.4	827	5	US-10-027-632-148684	Sequence 148684,
c 443	44	1.4	827	5	US-10-027-632-148685	Sequence 148685,
c 444	44	1.4	827	6	US-10-027-632-148684	Sequence 148684,
c 445	44	1.4	827	6	US-10-027-632-148685	Sequence 148685,
c 446	44	1.4	873	5	US-10-027-632-254541	Sequence 254541,
c 447	44	1.4	873	6	US-10-027-632-254541	Sequence 254541,
448	44	1.4	956	6	US-10-170-097-56	Sequence 56, Appl
449	44	1.4	956	8	US-10-926-684-56	Sequence 56, Appl
c 450	44	1.4	1005	4	US-09-925-065A-726951	Sequence 726951,
c 451	44	1.4	1005	4	US-09-925-065A-726952	Sequence 726952,
c 452	44	1.4	1051	5	US-10-027-632-31546	Sequence 31546, A
c 453	44	1.4	1051	5	US-10-027-632-31547	Sequence 31547, A
c 454	44	1.4	1051	5	US-10-027-632-31548	Sequence 31548, A
c 455	44	1.4	1051	6	US-10-027-632-31546	Sequence 31546, A
c 456	44	1.4	1051	6	US-10-027-632-31547	Sequence 31547, A
c 457	44	1.4	1051	6	US-10-027-632-31548	Sequence 31548, A
458	44	1.4	1068	4	US-09-925-065A-716304	Sequence 716304,
459	44	1.4	1068	4	US-09-925-065A-716305	Sequence 716305,
460	44	1.4	1068	4	US-09-925-065A-716306	Sequence 716306,
461	44	1.4	1068	4	US-09-925-065A-716307	Sequence 716307,
462	44	1.4	1068	4	US-09-925-065A-716308	Sequence 716308,
c 463	44	1.4	1103	5	US-10-027-632-257954	Sequence 257954,
c 464	44	1.4	1103	5	US-10-027-632-257955	Sequence 257955,
c 465	44	1.4	1103	6	US-10-027-632-257954	Sequence 257954,
c 466	44	1.4	1103	6	US-10-027-632-257955	Sequence 257955,
c 467	44	1.4	1122	4	US-09-925-065A-551840	Sequence 551840,
468	44	1.4	1240	5	US-10-027-632-263693	Sequence 263693,
469	44	1.4	1240	6	US-10-027-632-263693	Sequence 263693,
470	44	1.4	1242	4	US-09-925-065A-712795	Sequence 712795,
471	44	1.4	1330	4	US-09-925-065A-58322	Sequence 58322, A
472	44	1.4	1330	4	US-09-925-065A-58323	Sequence 58323, A
473	44	1.4	1330	4	US-09-925-065A-58324	Sequence 58324, A
c 474	44	1.4	1390	4	US-09-925-065A-47642	Sequence 47642, A
475	44	1.4	1427	4	US-09-925-065A-77993	Sequence 77993, A
476	44	1.4	1734	5	US-10-027-632-100291	Sequence 100291,
477	44	1.4	1734	6	US-10-027-632-100291	Sequence 100291,
478	44	1.4	1753	4	US-09-925-065A-714611	Sequence 714611,
479	44	1.4	1753	4	US-09-925-065A-714612	Sequence 714612,
480	44	1.4	1753	4	US-09-925-065A-714613	Sequence 714613,
481	44	1.4	1753	4	US-09-925-065A-714614	Sequence 714614,
c 482	44	1.4	1900	5	US-10-027-632-263112	Sequence 263112,
c 483	44	1.4	1900	6	US-10-027-632-263112	Sequence 263112,
484	44	1.4	1970	4	US-09-925-065A-721388	Sequence 721388,
c 485	44	1.4	2078	4	US-09-925-065A-16796	Sequence 16796, A
486	44	1.4	2184	4	US-09-925-065A-715885	Sequence 715885,
487	44	1.4	2184	4	US-09-925-065A-715886	Sequence 715886,
488	44	1.4	2184	4	US-09-925-065A-715887	Sequence 715887,
489	44	1.4	2184	4	US-09-925-065A-715888	Sequence 715888,
c 490	44	1.4	2725	4	US-09-925-065A-696477	Sequence 696477,
491	44	1.4	2845	9	US-10-956-157-2234	Sequence 2234, Ap
c 492	44	1.4	3488	4	US-09-925-065A-84868	Sequence 84868, A
493	44	1.4	5307	5	US-10-074-045-59	Sequence 59, Appl
494	44	1.4	5307	5	US-10-074-045-61	Sequence 61, Appl
c 495	44	1.4	6467	6	US-10-172-118-15	Sequence 15, Appl
c 496	44	1.4	6467	7	US-10-342-887-15	Sequence 15, Appl
c 497	44	1.4	6467	9	US-10-756-149-2274	Sequence 2274, Ap
c 498	44	1.4	7720	3	US-09-954-456-946	Sequence 946, App
c 499	44	1.4	7720	3	US-09-954-456-1589	Sequence 1589, Ap
c 500	44	1.4	7720	3	US-09-968-007A-125	Sequence 125, App

ALIGNMENTS

RESULT 1
US-10-450-763-29532
; Sequence 29532, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 29532
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (241)..(699)
; OTHER INFORMATION: 30% homologous to Leishmania major L8453.1,accession number
; OTHER INFORMATION: AC008054,Smith-Waterman Score=81.
US-10-450-763-29532

```
  Query Match            24.0%;  Score 748;  DB 9;  Length 850;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches  848;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        716 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGTAGCGCCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCG 60

Qy        776 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGATGGGGGTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCG 120

Qy        836 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCCAGCCGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCT 180

Qy        896 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTC 240

Qy        956 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAACCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTC 300

Qy       1016 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACCTGCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGC 360

Qy       1076 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAGTTCGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACAT 420

Qy       1136 GCGACGCTCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGT 1195
              ||||||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        421 GCGACGCGCGCTGGAGCTGGGCGCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGT 480

Qy       1196 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGCACAGGTGTGGCTGGCGCCTCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAG 540

Qy       1256 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTGCGGCTCGAGGCGGAGGGCGACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGA 600

Qy       1316 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTCCTTCAGGTGGGCGAGATGATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTG 660

Qy       1376 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACCGTGCAAGCCCGGCAGGCGGCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCC 720

Qy       1436 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCTCGGTCGTGTCCTTGCAGGAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGC 780

Qy       1496 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCCATCCTTTTCGGCGCCGTGCTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAA 840

Qy       1556 GCTGAGCTGA 1565
              ||||||||||
Db        841 GCTGAGCTGA 850


*RESULT 2
US-10-029-386-20699/c
; Sequence 20699, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20699
;   LENGTH: 708
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010615.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: AW302149.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P31735, EVALUE 5.50e+00
US-10-029-386-20699

  Query Match            21.0%;  Score 657;  DB 6;  Length 708;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  707;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        858 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 ATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGACGACTGCGTGCTAC 649

Qy        918 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 CACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACCTGCGGCAGGAGCTG 589

Qy        978 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 CAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTGCGCCCGGCTGACTGCT 529

Qy       1038 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 GTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTTCGAGCGGCTCTGGGTG 469

Qy       1098 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACGCTCGCTGGAGCTGGGC 409

Qy       1158 GCCGCGTTCCCGCTGCACGCGCCGCGGCGACCGCTGGTGCGCACAGGTGTGGCTGGCGCC 1217
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        408 GCCGCGTTCCCGCTGCACGCGCCGCGGCGGCCGCTGGTGCGCACAGGTGTGGCTGGCGCC 349

Qy       1218 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TCCTCCGGCGTGGCGGCGCGCGCGCTGAGCACCCGCAGCCTGCGGCTCGAGGCGGAGGGC 289

Qy       1278 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GACTTCGACGTCGCGGACCTGCGGGAGCTGGAGCGCGAGGTCCTTCAGGTGGGCGAGATG 229

Qy       1338 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATCGACAACATGGAGATGAAGGTCAACGTGCCCCGCTGGACCGTGCAAGCCCGGCAGGCG 169

Qy       1398 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GCGGGCGCCGAGCTCCTGTCCACGGTCAGCGCCGGCCCCTCCTCGGTCGTGTCCTTGCAG 109

Qy       1458 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GAGCGCGGGGGGGGGTTGCGACCCCAGGAAGGCCCTGGCCGCCATCCTTTTCGGCGCCGTG 49

Qy       1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         48 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAAGCTGAGCTGA 1
```

.

```
RESULT 3
US-10-029-386-6988/c
; Sequence 6988, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6988
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010615.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: AW302149.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q06805, EVALUE 1.90e-01
;   OTHER INFORMATION: NT HIT: gi4507086, EVALUE 7.80e+00
US-10-029-386-6988

  Query Match             16.8%;  Score 524;  DB 6;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-265;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        484 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 CTCTGCAGCCTGCTTGCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGC 465

Qy        544 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 TTCGCAGGGAGCCACCGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGT 405

Qy        604 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GCGCCTGGGGAGGATGGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCC 345

Qy        664 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CCGTCCGCAGAGGCGCACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCG 285

Qy        724 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CCGAGCGAGTCACGGACCATGAAGAGCGTTCGTGCCGCGCGGCCCAAGGCCGGGATGGGG 225

Qy        784 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCCGGGCCCAGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTTAGCCACATCCTGCCGCGCTGAGGGGGAGGCTAACGGGCGCGGGCGGCCCGGGCCCAGC 165

Qy        844 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CGGAGCCCACCGCGATGGCGAGGGAGGAGTGCAAGGCGCTGCTGGACGGGCTCAACAAGA 105

Qy        904 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAACC 45

Qy        964 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1007
              |||||||||||||||||||||||||||||||||||||||||||
```

```
Db         44 TGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTG 1

RESULT 4
US-09-925-065A-740956/c
; Sequence 740956, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 740956
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-740956

  Query Match             16.3%;  Score 509;  DB 4;  Length 560;
  Best Local Similarity   99.8%;  Pred. No. 5.3e-257;
  Matches  559;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        199 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 501

Qy        259 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 441

Qy        319 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 381

Qy        379 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAA 438
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAA 321

Qy        439 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 261

Qy        499 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 201

Qy        559 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 141

Qy        619 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 81

Qy        679 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 21

Qy        739 ACCATGAAGAGCGTTCGTGC 758
```

```
                    ||||||||||||||||||||
Db         20 ACCATGAAGAGCGTTCGTGC 1
```

RESULT 5
US-09-925-065A-769508/c
; Sequence 769508, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769508
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-769508

Query Match 16.2%; Score 506; DB 4; Length 557;
Best Local Similarity 99.8%; Pred. No. 2e-255;
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        202 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 261
              ||||||||||||||||||||
```

```
Qy        743 TGAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db         16 TGAAGAGCGTTCGTGC 1
```

RESULT 7
US-09-925-065A-737120
; Sequence 737120, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737120
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-737120

Query Match 16.1%; Score 504; DB 4; Length 555;
Best Local Similarity 99.8%; Pred. No. 2.3e-254;
Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        204 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 60

Qy        264 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 120

Qy        324 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 180

Qy        384 GATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 443
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 240

Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 300

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 360

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 420

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 480

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 540

Qy        744 GAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db        541 GAAGAGCGTTCGTGC 555
```

RESULT 8
US-10-450-763-8312/c
; Sequence 8312, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 8312
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (17)..(1054)
; OTHER INFORMATION: 30% homologous to Streptomyces coelicolor A3(2) putative
; OTHER INFORMATION: translation initiation factor IF-2(fragment),accession number
; OTHER INFORMATION: AL355913,Smith-Waterman Score=196.
US-10-450-763-8312

Query Match 7.7%; Score 241; DB 9; Length 1349;
Best Local Similarity 100.0%; Pred. No. 1.1e-115;
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        902 GACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAA 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GACGACTGCGTGCTACCACCACCTGGTGCTGACCGTCGGTGGCTCGGCGGACTCGCAGAA 343

Qy        962 CCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTG 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CCTGCGGCAGGAGCTGCAAAAGACGCGCCAGAAGGCGCAGGAGCTGGCGGTGTCCACCTG 283

Qy       1022 CGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTT 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCCTGGCCGCCGACGAGCGCGCCGAGTT 223

Qy       1082 CGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACG 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGAGCGGCTCTGGGTGGCCTTCTCGGGCTGCCTGGACCTGCTGGAAGCGGACATGCGACG 163

Qy       1142 C 1142
              |
Db        162 C 162
```

RESULT 9
US-09-908-975-23330
; Sequence 23330, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE

```
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23330
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-23330

  Query Match             1.9%;  Score 60;  DB 3;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2464 GGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGG 2523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCTGGGAATCCTCTTTAGAGCACTTAATCCTATTTATCCCCTGGAATGTGCGTGCTGG 60


RESULT 10
US-10-723-860-1753/c
; Sequence 1753, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1753
;   LENGTH: 175737
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-1753

  Query Match             1.7%;  Score 53;  DB 8;  Length 175737;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47829 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 47777


RESULT 11
US-10-783-271-34/c
; Sequence 34, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
;   LENGTH: 175737
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-34

  Query Match             1.7%;  Score 53;  DB 9;  Length 175737;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47829 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 47777


RESULT 12
US-10-674-124A-76/c
; Sequence 76, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 76
;   LENGTH: 428
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AL357552.12_93957
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 1
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 8481068
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 60131
US-10-674-124A-76

  Query Match             1.7%;  Score 52;  DB 8;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 116


RESULT 13
US-09-925-065A-591947/c
; Sequence 591947, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
```

```
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591947
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591947

  Query Match             1.7%;  Score 52;  DB 4;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 113


RESULT 14
US-10-027-632-41688
; Sequence 41688, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41688
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-41688

  Query Match             1.7%;  Score 52;  DB 5;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 379


RESULT 15
US-10-027-632-41688
; Sequence 41688, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41688
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-41688

  Query Match             1.7%;  Score 52;  DB 6;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 379


RESULT 16
US-09-925-065A-768118
; Sequence 768118, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 768118
;   LENGTH: 574
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-768118

  Query Match             1.7%;  Score 52;  DB 4;  Length 574;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        235 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 286


RESULT 17
US-09-925-065A-59682/c
; Sequence 59682, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59682
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-59682

  Query Match             1.7%;  Score 52;  DB 4;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 124


RESULT 18
US-09-925-065A-386361/c
; Sequence 386361, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 386361
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-386361

  Query Match             1.7%;  Score 52;  DB 4;  Length 617;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255


RESULT 19
US-10-027-632-61987
; Sequence 61987, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61987
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-61987

  Query Match             1.7%;  Score 52;  DB 5;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 20
US-10-027-632-61988
; Sequence 61988, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
```

```
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61988
;    LENGTH: 631
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-61988

  Query Match             1.7%;  Score 52;  DB 5;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 21
US-10-027-632-61989
; Sequence 61989, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61989
;    LENGTH: 631
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-61989

  Query Match             1.7%;  Score 52;  DB 5;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 22
US-10-027-632-61987
; Sequence 61987, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61987
;    LENGTH: 631
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-61987

  Query Match             1.7%;  Score 52;  DB 6;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 23
US-10-027-632-61988
; Sequence 61988, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61988
;    LENGTH: 631
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-61988

  Query Match             1.7%;  Score 52;  DB 6;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 24
US-10-027-632-61989
; Sequence 61989, Application US/10027632
```

```
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61989
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-61989

  Query Match             1.7%;  Score 52;  DB 6;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 5e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 412


RESULT 25
US-09-925-065A-708701/c
; Sequence 708701, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 708701
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-708701

  Query Match             1.7%;  Score 52;  DB 4;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 439


RESULT 26
US-10-450-763-14141/c
; Sequence 14141, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 14141
;   LENGTH: 1437
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (253)..(786)
;   OTHER INFORMATION: 50% homologous to Homo sapiens capping protein alpha subunit
;   OTHER INFORMATION: isoform 1,accession number U56637,Smith-Waterman Score=380.
US-10-450-763-14141

  Query Match             1.7%;  Score 52;  DB 9;  Length 1437;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 192


RESULT 27
US-09-764-891-8895/c
; Sequence 8895, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8895
;   LENGTH: 9620
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-8895

  Query Match             1.7%;  Score 52;  DB 3;  Length 9620;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7058 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 7007


RESULT 28
US-09-918-686-1
; Sequence 1, Application US/09918686
```

```
; Patent No. US20020076720A1
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary
;  APPLICANT:  Proll, Sean
;  APPLICANT:  Paeper, Bryan
;  APPLICANT:  Staehling-Hampton, Karen
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING
;  TITLE OF INVENTION:  GENOMIC DELETIONS
;  FILE REFERENCE: 240083.515
;  CURRENT APPLICATION NUMBER: US/09/918,686
;  CURRENT FILING DATE:  2001-07-30
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 92139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 7043, 8369, 8401
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

  Query Match             1.7%;  Score 52;  DB 3;  Length 92139;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85294 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85345


RESULT 29
US-10-353-150-1
; Sequence 1, Application US/10353150
; Publication No. US20030157543A1
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary E.
;  APPLICANT:  Proll, Sean
;  APPLICANT:  Paeper, Bryan
;  APPLICANT:  Staehling-Hampton, Karen
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING
;  TITLE OF INVENTION:  GENOMIC DELETIONS
;  FILE REFERENCE: 240083.515C1
;  CURRENT APPLICATION NUMBER: US/10/353,150
;  CURRENT FILING DATE:  2003-01-27
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 92139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 7043, 8369, 8401
;   OTHER INFORMATION: n = A,T,C or G
US-10-353-150-1

  Query Match             1.7%;  Score 52;  DB 6;  Length 92139;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85294 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 85345


RESULT 30
US-10-408-168-1/c
; Sequence 1, Application US/10408168
; Publication No. US20030235847A1
; GENERAL INFORMATION:
;  APPLICANT: Paeper, Bryan W.
;  APPLICANT:  Proll, Sean
;  APPLICANT:  Charmley, Patrick R.
;  APPLICANT:  Brunkow, Mary E.
;  APPLICANT:  Uitterlinden, Andreas Gerardus
;  TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
;  TITLE OF INVENTION:  GENE REGION WITH BONE MINERAL DENSITY
;  FILE REFERENCE: 240083.525
;  CURRENT APPLICATION NUMBER: US/10/408,168
;  CURRENT FILING DATE:  2003-04-03
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 130320
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 100722, 100754, 102080, 117731, 124408, 124532, 124585, 124955,
;   LOCATION: 124956, 124963, 124964, 124965, 124967
;   OTHER INFORMATION: n = A,T,C or G
US-10-408-168-1

  Query Match             1.7%;  Score 52;  DB 6;  Length 130320;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23819 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 23768


RESULT 31
US-10-085-117-244
; Sequence 244, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: 529452000121
;  CURRENT APPLICATION NUMBER: US/10/085,117
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244
;   LENGTH: 136726
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: (1)...(136726)
;   OTHER INFORMATION: n = any nucleotide
US-10-085-117-244

  Query Match             1.7%;  Score 52;  DB 6;  Length 136726;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66522 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 66573


RESULT 32
US-10-981-277-50
; Sequence 50, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa
;  TITLE OF INVENTION: Compositions and Methods for Glioma Classification
```

; FILE REFERENCE: 03-968-US
; CURRENT APPLICATION NUMBER: US/10/981,277
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/516,817
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 50
; LENGTH: 150437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-981-277-50

Query Match 1.7%; Score 52; DB 9; Length 150437;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125751 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 125802

RESULT 33
US-10-174-014-12
; Sequence 12, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PTS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 12
; LENGTH: 221000
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77967
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77968
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77969
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77970
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77971
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77972
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77973
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77974
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77975
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77976
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77977
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77978
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77979
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77980
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77981
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77982
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77983
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77984
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77985
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77986
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77987
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77988
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77989
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77990
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77991
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77992
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77993

```
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77994
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77995
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77996
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77997
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77998
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 77999
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78000
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78001
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78002
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78003
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78004
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78005
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78006
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78007
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78008
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78009
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78010
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78011
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 78012
;   OTHER INFORMATION: unknown
;   FEATURE:

  Query Match             1.7%;  Score 52;  DB 6;  Length 221000;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173223 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 173274


RESULT 34
US-10-087-192-652
; Sequence 652, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 652
;   LENGTH: 233380
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(233380)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-652

  Query Match             1.7%;  Score 52;  DB 5;  Length 233380;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     179881 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 179932


RESULT 35
US-10-741-601-5719/c
; Sequence 5719, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5719
;   LENGTH: 398287
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(398287)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
```

US-10-741-601-5719

Query Match 1.7%; Score 52; DB 7; Length 398287;
Best Local Similarity 100.0%; Pred. No. 4.2e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 281811 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 281760

RESULT 36
US-10-741-600-17839/c
; Sequence 17839, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17839
; LENGTH: 398287
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(398287)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17839

Query Match 1.7%; Score 52; DB 8; Length 398287;
Best Local Similarity 100.0%; Pred. No. 4.2e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 281811 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 281760

RESULT 37
US-09-925-065A-211249/c
; Sequence 211249, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 211249
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-211249

Query Match 1.6%; Score 51; DB 4; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 167 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 117

RESULT 38
US-09-925-065A-141833
; Sequence 141833, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 141833
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-141833

Query Match 1.6%; Score 51; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 402 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 452

RESULT 39
US-10-027-632-54553
; Sequence 54553, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 54553
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-54553

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 40
US-10-027-632-55034
; Sequence 55034, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55034
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-55034

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 41
US-10-027-632-56282
; Sequence 56282, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56282
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-56282

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 42
US-10-027-632-56384
; Sequence 56384, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56384
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-56384

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 43
US-10-027-632-180536
; Sequence 180536, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180536
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-180536

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 44
US-10-027-632-314485
; Sequence 314485, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 314485
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-314485

  Query Match             1.6%;  Score 51;  DB 5;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 45
US-10-027-632-54553
; Sequence 54553, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54553
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-54553

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 46
US-10-027-632-55034
; Sequence 55034, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
```

```
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55034
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-55034

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 47
US-10-027-632-56282
; Sequence 56282, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56282
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-56282

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 48
US-10-027-632-56384
; Sequence 56384, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56384
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-56384

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 49
US-10-027-632-180536
; Sequence 180536, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180536
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-180536

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 50
US-10-027-632-314485
; Sequence 314485, Application US/10027632
```

```
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 314485
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-314485

  Query Match             1.6%;  Score 51;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 451


RESULT 51
US-09-925-298-220/c
; Sequence 220, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103
;  CURRENT APPLICATION NUMBER: US/09/925,298
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
;   LENGTH: 832
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-298-220

  Query Match             1.6%;  Score 51;  DB 3;  Length 832;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 26


RESULT 52
US-10-102-806-220/c
; Sequence 220, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
;   LENGTH: 832
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-102-806-220

  Query Match             1.6%;  Score 51;  DB 5;  Length 832;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 26


RESULT 53
US-10-450-763-29534
; Sequence 29534, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 29534
;   LENGTH: 2791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (18)..(194)
;   OTHER INFORMATION: 43% homologous to Trypanoplasma borreli ribosomal protein
;   OTHER INFORMATION: S12,accession number U14183,Smith-Waterman Score=95.
US-10-450-763-29534

  Query Match             1.6%;  Score 51;  DB 9;  Length 2791;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        465 CCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACC 515
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1538 CCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCCGGAGTTGGCACC 1588


RESULT 54
US-10-737-082-22/c
; Sequence 22, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Bayer Healthcare LLC
;  APPLICANT:  Beard, Chris
;  APPLICANT:  Burgess, Chris
;  APPLICANT:  Gannon, Allison
;  APPLICANT:  Harvey, Jeanne
;  APPLICANT:  Lechner, John F.
;  APPLICANT:  Li, Zheng
;  TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
;  FILE REFERENCE: 1657/2032
;  CURRENT APPLICATION NUMBER: US/10/737,082
;  CURRENT FILING DATE:  2003-12-16
;  PRIOR APPLICATION NUMBER: US 10/737,082
;  PRIOR FILING DATE: 2003-12-16
;  NUMBER OF SEQ ID NOS: 300
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
;   LENGTH: 23139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-737-082-22

  Query Match             1.6%;  Score 51;  DB 9;  Length 23139;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 3733


RESULT 55
US-10-765-790-22/c
; Sequence 22, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
;  APPLICANT: Bayer Healthcare LLC
;  APPLICANT:  Beard, Chris
;  APPLICANT:  Burgess, Chris
;  APPLICANT:  Gannon, Allison
;  APPLICANT:  Harvey, Jeanne
;  APPLICANT:  Lechner, John F.
;  APPLICANT:  Li, Zheng
;  TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
;  FILE REFERENCE: 1657/2035
;  CURRENT APPLICATION NUMBER: US/10/765,790
;  CURRENT FILING DATE:  2004-01-27
;  PRIOR APPLICATION NUMBER: US 10/737,082
;  PRIOR FILING DATE: 2003-12-16
;  NUMBER OF SEQ ID NOS: 300
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
;   LENGTH: 23139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-765-790-22

  Query Match             1.6%;  Score 51;  DB 9;  Length 23139;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 3733


RESULT 56
US-10-840-590-3/c
; Sequence 3, Application US/10840590
; Publication No. US20050233341A1
; GENERAL INFORMATION:
;  APPLICANT: ROTH, RICHARD B.
;  APPLICANT:  NELSON, MATTHEW ROBERTS
;  APPLICANT:  KAMMERER, STEFAN M.
;  APPLICANT:  BRAUN, ANDREAS
;  APPLICANT:  HOYAL-WRIGHTSON, CAROLYN R.
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF MELANOMA AND TREATMENTS THEREOF
;  FILE REFERENCE: SEQ-4061-CP
;  CURRENT APPLICATION NUMBER: US/10/840,590
;  CURRENT FILING DATE:  2004-05-05
;  PRIOR APPLICATION NUMBER: 60/489,703
;  PRIOR FILING DATE: 2003-07-23
;  NUMBER OF SEQ ID NOS: 1638
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 68200
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (50036)..(50036)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-840-590-3

  Query Match             1.6%;  Score 51;  DB 9;  Length 68200;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63277 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 63227


RESULT 57
US-10-087-192-1360/c
; Sequence 1360, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1360
;   LENGTH: 215221
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-1360

  Query Match             1.6%;  Score 51;  DB 5;  Length 215221;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-15;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50858 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 50808


RESULT 58
US-09-973-278-870
; Sequence 870, Application US/09973278
; Publication No. US20040044191A1
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P2
;  CURRENT APPLICATION NUMBER: US/09/973,278
;  CURRENT FILING DATE:  2001-10-10
;  PRIOR APPLICATION NUMBER: 60/239,899
```

```
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 09/227,357
;  PRIOR FILING DATE: 1999-01-08
;  PRIOR APPLICATION NUMBER: PCT/US98/13684
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/051,926
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,793
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,925
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,929
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,803
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,732
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,931
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,932
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,916
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,930
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,918
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,920
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,733
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/052,795
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,919
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/051,928
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: 60/055,722
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,723
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,948
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,949
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,953
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,950
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,947
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,964
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/056,360
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,684
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,984
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,954
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/058,785
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,664
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,660
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,661
;  PRIOR FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 947
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 870
;    LENGTH: 288
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-973-278-870

  Query Match             1.6%;  Score 50;  DB 3;  Length 288;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 270


RESULT 59
US-10-450-763-29529
; Sequence 29529, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 29529
;    LENGTH: 432
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: SIMILAR
;    LOCATION: (413)..(120)
;    OTHER INFORMATION: 33% homologous to Human herpesvirus 6 H87,accession number
;    OTHER INFORMATION: U92288,Smith-Waterman Score=105.
US-10-450-763-29529

  Query Match             1.6%;  Score 50;  DB 9;  Length 432;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        888 GACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGT 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GACGGGCTCAACAAGACGACTGCGTGCTACCACCACCTGGTGCTGACCGT 71


RESULT 60
US-09-925-065A-484883/c
; Sequence 484883, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 484883
;   LENGTH: 571
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-484883

  Query Match             1.6%;  Score 50;  DB 4;  Length 571;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 222


RESULT 61
US-09-925-065A-785967/c
; Sequence 785967, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785967
;   LENGTH: 572
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785967

  Query Match             1.6%;  Score 50;  DB 4;  Length 572;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 298


RESULT 62
US-09-925-065A-894777/c
; Sequence 894777, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 894777
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-894777

  Query Match             1.6%;  Score 50;  DB 4;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 180


RESULT 63
US-09-925-065A-892664/c
; Sequence 892664, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892664
;   LENGTH: 630
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892664

  Query Match             1.6%;  Score 50;  DB 4;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 185


RESULT 64
US-09-925-065A-892405/c
; Sequence 892405, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892405
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892405

  Query Match             1.6%;  Score 50;  DB 4;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 188


RESULT 65
US-09-925-065A-892514/c
; Sequence 892514, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892514
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892514

  Query Match             1.6%;  Score 50;  DB 4;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 188


RESULT 66
US-09-925-065A-949927
; Sequence 949927, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 949927
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-949927

  Query Match             1.6%;  Score 50;  DB 4;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 78


RESULT 67
US-09-925-065A-917225
; Sequence 917225, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 917225
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-917225

  Query Match             1.6%;  Score 50;  DB 4;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497


RESULT 68
US-09-925-065A-917226
; Sequence 917226, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 917226
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-917226

  Query Match             1.6%;  Score 50;  DB 4;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497


RESULT 69
US-09-925-065A-944569
; Sequence 944569, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944569
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944569

  Query Match             1.6%;  Score 50;  DB 4;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497


RESULT 70
US-10-027-632-158850/c
; Sequence 158850, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 158850
;   LENGTH: 714
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-158850

  Query Match             1.6%;  Score 50;  DB 5;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 84


RESULT 71
US-10-027-632-158850/c
; Sequence 158850, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 158850
;   LENGTH: 714
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-158850

  Query Match             1.6%;  Score 50;  DB 6;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 84
```

RESULT 72
US-09-925-065A-934608
; Sequence 934608, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 934608
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-934608

Query Match 1.6%; Score 50; DB 4; Length 739;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 652
```

RESULT 73
US-09-925-065A-941367
; Sequence 941367, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941367
; LENGTH: 742
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-941367

Query Match 1.6%; Score 50; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 652
```

RESULT 74
US-10-027-632-265269
; Sequence 265269, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 265269
; LENGTH: 1986
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-265269

Query Match 1.6%; Score 50; DB 5; Length 1986;
Best Local Similarity 100.0%; Pred. No. 5.5e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 246
```

RESULT 75
US-10-027-632-265269
; Sequence 265269, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

```
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 265269
;    LENGTH: 1986
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-265269

  Query Match             1.6%;  Score 50;  DB 6;  Length 1986;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 246


RESULT 76
US-10-374-979-10
; Sequence 10, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
;  APPLICANT: John P. Carulli et al.
;  TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
;  FILE REFERENCE: 032796-021
;  CURRENT APPLICATION NUMBER: US/10/374,979
;  CURRENT FILING DATE:  2003-03-04
;  PRIOR APPLICATION NUMBER: US 09/544,398
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: US 09/543,771
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: US 09/229,319
;  PRIOR FILING DATE: 1999-01-13
;  PRIOR APPLICATION NUMBER: US 60/071,449
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/105,511
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 10
;    LENGTH: 8705
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-374-979-10

  Query Match             1.6%;  Score 50;  DB 6;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 77
US-10-182-936A-10
; Sequence 10, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Kristina M.
;  APPLICANT:  Anisowicz, Anthony
;  APPLICANT:  Bhat, Bheem
;  APPLICANT:  Damagnez, Veronique
;  APPLICANT:  Robinson, John
;  APPLICANT:  Yaworsky, Paul
;  TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
;  FILE REFERENCE: 032796-143
;  CURRENT APPLICATION NUMBER: US/10/182,936A
;  CURRENT FILING DATE:  2002-08-02
;  PRIOR APPLICATION NUMBER: PCT/US02/15982
;  PRIOR FILING DATE: 2002-05-17
;  PRIOR APPLICATION NUMBER: US 60/291,311
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: US 60/353,058
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: US 60/361,293
;  PRIOR FILING DATE: 2002-03-04
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;    LENGTH: 8705
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-182-936A-10

  Query Match             1.6%;  Score 50;  DB 7;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 78
US-10-731-739-10
; Sequence 10, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
;  APPLICANT: Carulli, John P.
;  APPLICANT:  Little, Randall D.
;  APPLICANT:  Recker, Robert R.
;  APPLICANT:  Johnson, Mark L.
;  TITLE OF INVENTION: High bone mass gene of 11q13.3
;  FILE REFERENCE: 032796-013
;  CURRENT APPLICATION NUMBER: US/10/731,739
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: US/09/544,398B
;  PRIOR FILING DATE: 2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/229,319
;  PRIOR FILING DATE: 1999-01-13
;  PRIOR APPLICATION NUMBER: US 60/071,449
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/105,511
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 641
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;    LENGTH: 8705
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-731-739-10

  Query Match             1.6%;  Score 50;  DB 8;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 79
US-10-477-238A-10
; Sequence 10, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
;  APPLICANT: Babij, Philip
;  APPLICANT:  Yaworsky, Paul
;  APPLICANT:  Bex, Frederick J. III
;  APPLICANT:  Bodine, Peter Van Nest
;  TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
;  FILE REFERENCE: 032796-212
;  CURRENT APPLICATION NUMBER: US/10/477,238A
;  CURRENT FILING DATE:  2003-11-10
;  PRIOR APPLICATION NUMBER: US 60/290,071
```

```
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/291,311
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: US 60/353,058
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: US 60/361,293
;  PRIOR FILING DATE: 2002-03-04
;  NUMBER OF SEQ ID NOS: 812
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-477-238A-10

  Query Match             1.6%;  Score 50;  DB 8;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 80
US-10-680-287A-10
; Sequence 10, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
;  APPLICANT: Babij, Philip
;  APPLICANT:  Yaworsky, Paul
;  APPLICANT:  Bex, Frederick J. III
;  APPLICANT:  Bodine, Peter Van Nest
;  TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
;  FILE REFERENCE: 032796-179
;  CURRENT APPLICATION NUMBER: US/10/680,287A
;  CURRENT FILING DATE:  2003-10-08
;  PRIOR APPLICATION NUMBER: PCT/US02/14876
;  PRIOR FILING DATE: 2002-05-13
;  PRIOR APPLICATION NUMBER: US 60/290,071
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/291,311
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: US 60/353,058
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: US 60/361,293
;  PRIOR FILING DATE: 2002-03-04
;  NUMBER OF SEQ ID NOS: 812
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-680-287A-10

  Query Match             1.6%;  Score 50;  DB 8;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 81
US-10-477-173-10
; Sequence 10, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
;  APPLICANT: Genome Therapeutics Corporation and
;  APPLICANT:  Allen, Kristina M.
;  APPLICANT:  Yaworsky, Paul
;  APPLICANT:  Morales, Arturo J.
;  APPLICANT:  Graham, James R.
;  APPLICANT:  Anisowicz, Anthony
;  APPLICANT:  Liu, Wei
;  TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
;  FILE REFERENCE: 032796-135
;  CURRENT APPLICATION NUMBER: US/10/477,173
;  CURRENT FILING DATE:  2003-11-10
;  PRIOR APPLICATION NUMBER: US 60/290,071
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/291,311
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: US 60/353,058
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: US 60/361,293
;  PRIOR FILING DATE: 2002-03-04
;  NUMBER OF SEQ ID NOS: 1086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-477-173-10

  Query Match             1.6%;  Score 50;  DB 9;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 82
US-10-834-377-10
; Sequence 10, Application US/10834377
; Publication No. US20050142617A1
; GENERAL INFORMATION:
;  APPLICANT: Carulli, John P.
;  APPLICANT:  Little, Randall D.
;  APPLICANT:  Recker, Robert R.
;  APPLICANT:  Johnson, Mark L.
;  TITLE OF INVENTION: High bone mass gene of 11q13.3
;  FILE REFERENCE: 032796-014
;  CURRENT APPLICATION NUMBER: US/10/834,377
;  CURRENT FILING DATE:  2004-04-29
;  PRIOR APPLICATION NUMBER: US/09/543,771B
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: US 09/229,319
;  PRIOR FILING DATE: 1999-01-13
;  PRIOR APPLICATION NUMBER: US 60/071,449
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/105,511
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 641
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 8705
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-834-377-10

  Query Match             1.6%;  Score 50;  DB 9;  Length 8705;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541


RESULT 83
US-09-764-891-8149/c
; Sequence 8149, Application US/09764891
```

```
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8149
;   LENGTH: 31474
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-8149

  Query Match             1.6%;  Score 50;  DB 3;  Length 31474;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 23


RESULT 84
US-09-764-878-379/c
; Sequence 379, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 379
;   LENGTH: 32189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-379

  Query Match             1.6%;  Score 50;  DB 3;  Length 32189;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12517 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12468


RESULT 85
US-10-079-854-379/c
; Sequence 379, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 379
;   LENGTH: 32189
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-079-854-379

  Query Match             1.6%;  Score 50;  DB 5;  Length 32189;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12517 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12468


RESULT 86
US-09-764-877-2623/c
; Sequence 2623, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2623
;   LENGTH: 32193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-2623

  Query Match             1.6%;  Score 50;  DB 3;  Length 32193;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28855 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 28806


RESULT 87
US-10-216-464-37/c
; Sequence 37, Application US/10216464
; Publication No. US20030207285A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ19C1
;  CURRENT APPLICATION NUMBER: US/10/216,464
;  CURRENT FILING DATE:  2002-08-12
;  PRIOR APPLICATION NUMBER: 09/764,883
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
;   LENGTH: 32193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-10-216-464-37

  Query Match             1.6%;  Score 50;  DB 6;  Length 32193;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28855 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 28806


RESULT 88
US-10-242-515-2623/c
; Sequence 2623, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005C1
;  CURRENT APPLICATION NUMBER: US/10/242,515
;  CURRENT FILING DATE:  2002-09-13
;  PRIOR APPLICATION NUMBER: 09/764,877
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2623
;   LENGTH: 32193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-242-515-2623

  Query Match             1.6%;  Score 50;  DB 6;  Length 32193;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28855 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 28806


RESULT 89
US-09-764-878-377/c
; Sequence 377, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 377
;   LENGTH: 32221
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (7464)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-878-377

  Query Match             1.6%;  Score 50;  DB 3;  Length 32221;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12548 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12499


RESULT 90
US-10-079-854-377/c
; Sequence 377, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 377
;   LENGTH: 32221
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7464)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-079-854-377

  Query Match             1.6%;  Score 50;  DB 5;  Length 32221;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12548 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 12499


RESULT 91
US-10-085-117-190
; Sequence 190, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: 529452000121
;  CURRENT APPLICATION NUMBER: US/10/085,117
;  CURRENT FILING DATE:  2002-02-27
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 190
;   LENGTH: 33126
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-117-190

  Query Match             1.6%;  Score 50;  DB 6;  Length 33126;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-15;
```

```
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30352 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 30401


RESULT 92
US-10-741-600-17637
; Sequence 17637, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17637
;   LENGTH: 56510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-17637

  Query Match             1.6%;  Score 50;  DB 8;  Length 56510;
  Best Local Similarity   100.0%;  Pred. No. 5e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAAC 2938
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4790 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAAC 4839


RESULT 93
US-10-087-192-574/c
; Sequence 574, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 574
;   LENGTH: 66973
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(66973)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-574

  Query Match             1.6%;  Score 50;  DB 5;  Length 66973;
  Best Local Similarity   100.0%;  Pred. No. 5e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29006 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 28957


RESULT 94
US-10-087-192-1408
; Sequence 1408, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1408
;   LENGTH: 156843
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(156843)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1408

  Query Match             1.6%;  Score 50;  DB 5;  Length 156843;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     148232 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 148281


RESULT 95
US-10-719-993-6812
; Sequence 6812, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6812
;   LENGTH: 202814
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-6812

  Query Match             1.6%;  Score 50;  DB 8;  Length 202814;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145090 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 145139


RESULT 96
US-10-271-416-9
; Sequence 9, Application US/10271416
; Publication No. US20040043021A1
; GENERAL INFORMATION:
;  APPLICANT: Keith, Tim
;  APPLICANT:  Little, Randall D.
;  APPLICANT:  Van Eerdewegh, Paul
;  APPLICANT:  Dupuis, Josee
```

```
;  APPLICANT:  Del Mastro, Richard G.
;  APPLICANT:  Allen, Kristina
;  TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
;  TITLE OF INVENTION:  RELATING TO RESPIRATORY DISEASES AND OBESITY
;  FILE REFERENCE: 2976-4045
;  CURRENT APPLICATION NUMBER: US/10/271,416
;  CURRENT FILING DATE:  2002-10-11
;  PRIOR APPLICATION NUMBER: 60/328,424
;  PRIOR FILING DATE: 2001-10-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 276820
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: conflict
;   LOCATION: (167043)...(167043)
;   OTHER INFORMATION: Y=C or T
US-10-271-416-9

  Query Match             1.6%;  Score 50;  DB 7;  Length 276820;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194263 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 194312


RESULT 97
US-09-925-065A-140670/c
; Sequence 140670, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 140670
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-140670

  Query Match             1.6%;  Score 49;  DB 4;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225


RESULT 98
US-09-925-065A-140671/c
; Sequence 140671, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 140671
;   LENGTH: 532
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-140671

  Query Match             1.6%;  Score 49;  DB 4;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 184


RESULT 99
US-09-925-065A-18609/c
; Sequence 18609, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18609
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18609

  Query Match             1.6%;  Score 49;  DB 4;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 100
US-09-925-065A-18610/c
```

```
; Sequence 18610, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18610
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18610

  Query Match             1.6%;  Score 49;  DB 4;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 101
US-09-925-065A-18611/c
; Sequence 18611, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18611
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18611

  Query Match             1.6%;  Score 49;  DB 4;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 102
US-10-027-632-80386/c
; Sequence 80386, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80386
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(580)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80386

  Query Match             1.6%;  Score 49;  DB 5;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 233


RESULT 103
US-10-027-632-80387/c
; Sequence 80387, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80387
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(580)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80387

  Query Match             1.6%;  Score 49;  DB 5;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 233


RESULT 104
US-10-027-632-80386/c
; Sequence 80386, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80386
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(580)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80386

  Query Match             1.6%;  Score 49;  DB 6;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 233


RESULT 105
US-10-027-632-80387/c
; Sequence 80387, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80387
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(580)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80387

  Query Match             1.6%;  Score 49;  DB 6;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 233


RESULT 106
US-09-925-065A-708700/c
; Sequence 708700, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 708700
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-708700

  Query Match             1.6%;  Score 49;  DB 4;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2892 GCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 439


RESULT 107
US-09-925-065A-932035/c
; Sequence 932035, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932035
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932035

  Query Match             1.6%;  Score 49;  DB 4;  Length 750;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 564


RESULT 108
US-09-925-065A-932036/c
; Sequence 932036, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932036
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932036

  Query Match             1.6%;  Score 49;  DB 4;  Length 750;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 564


RESULT 109
US-10-027-632-113573
; Sequence 113573, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113573
;   LENGTH: 3170
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-113573

  Query Match             1.6%;  Score 49;  DB 5;  Length 3170;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 2033


RESULT 110
US-10-027-632-113574
; Sequence 113574, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
```

```
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113574
;   LENGTH: 3170
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-113574

  Query Match             1.6%;  Score 49;  DB 5;  Length 3170;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 2033


RESULT 111
US-10-027-632-113573
; Sequence 113573, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113573
;   LENGTH: 3170
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-113573

  Query Match             1.6%;  Score 49;  DB 6;  Length 3170;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 2033


RESULT 112
US-10-027-632-113574
; Sequence 113574, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113574
;   LENGTH: 3170
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-113574

  Query Match             1.6%;  Score 49;  DB 6;  Length 3170;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-14;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1985 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 2033


RESULT 113
US-10-027-632-182264
; Sequence 182264, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182264
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-182264

  Query Match             1.5%;  Score 48;  DB 5;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207


RESULT 114
US-10-027-632-319423
```

```
; Sequence 319423, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 319423
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-319423

  Query Match             1.5%;  Score 48;  DB 5;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207


RESULT 115
US-10-027-632-182264
; Sequence 182264, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182264
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-182264

  Query Match             1.5%;  Score 48;  DB 6;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207


RESULT 116
US-10-027-632-319423
; Sequence 319423, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 319423
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-319423

  Query Match             1.5%;  Score 48;  DB 6;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207


RESULT 117
US-10-027-632-132626
; Sequence 132626, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
```

```
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132626
;   LENGTH: 716
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-132626

  Query Match             1.5%;  Score 48;  DB 5;  Length 716;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 405


RESULT 118
US-10-027-632-132626
; Sequence 132626, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 132626
;   LENGTH: 716
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-132626

  Query Match             1.5%;  Score 48;  DB 6;  Length 716;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 405


RESULT 119
US-10-052-482-172
; Sequence 172, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
;  APPLICANT: Engelhard, Eric
;  APPLICANT:  Morris, David
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71087/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/10/052,482
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 172
;   LENGTH: 52242
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51242)..(51261)
;   OTHER INFORMATION: "n" at positions 51242 to 51261 can be any base
US-10-052-482-172

  Query Match             1.5%;  Score 48;  DB 7;  Length 52242;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5520 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 5567


RESULT 120
US-10-737-082-85/c
; Sequence 85, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
;  APPLICANT: Bayer Healthcare LLC
;  APPLICANT:  Beard, Chris
;  APPLICANT:  Burgess, Chris
;  APPLICANT:  Gannon, Allison
;  APPLICANT:  Harvey, Jeanne
;  APPLICANT:  Lechner, John F.
;  APPLICANT:  Li, Zheng
;  TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
;  FILE REFERENCE: 1657/2032
;  CURRENT APPLICATION NUMBER: US/10/737,082
;  CURRENT FILING DATE:  2003-12-16
;  PRIOR APPLICATION NUMBER: US 10/737,082
;  PRIOR FILING DATE: 2003-12-16
;  NUMBER OF SEQ ID NOS: 300
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 85
;   LENGTH: 53779
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-737-082-85

  Query Match             1.5%;  Score 48;  DB 9;  Length 53779;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCA 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2311 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCA 2264


RESULT 121
US-10-765-790-85/c
; Sequence 85, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
;  APPLICANT: Bayer Healthcare LLC
;  APPLICANT:  Beard, Chris
;  APPLICANT:  Burgess, Chris
;  APPLICANT:  Gannon, Allison
;  APPLICANT:  Harvey, Jeanne
;  APPLICANT:  Lechner, John F.
;  APPLICANT:  Li, Zheng
;  TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
```

```
;  FILE REFERENCE: 1657/2035
;  CURRENT APPLICATION NUMBER: US/10/765,790
;  CURRENT FILING DATE:  2004-01-27
;  PRIOR APPLICATION NUMBER: US 10/737,082
;  PRIOR FILING DATE: 2003-12-16
;  NUMBER OF SEQ ID NOS: 300
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 85
;   LENGTH: 53779
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-765-790-85

  Query Match             1.5%;  Score 48;  DB 9;  Length 53779;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCA 2936
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2311 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCA 2264


RESULT 122
US-10-741-600-17651
; Sequence 17651, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17651
;   LENGTH: 141121
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-17651

  Query Match             1.5%;  Score 48;  DB 8;  Length 141121;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       8800 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 8847


RESULT 123
US-10-087-192-790
; Sequence 790, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790
;   LENGTH: 181684
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(181684)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-790

  Query Match             1.5%;  Score 48;  DB 5;  Length 181684;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-14;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db     155726 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 155773


RESULT 124
US-10-357-930-16322
; Sequence 16322, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16322
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-16322

  Query Match             1.5%;  Score 47;  DB 8;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3070 GCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 424


RESULT 125
US-10-357-930-46148
; Sequence 46148, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
```

```
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46148
;   LENGTH: 516
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 495, 496
;   OTHER INFORMATION: n = A,T,C or G
US-10-357-930-46148

  Query Match             1.5%;  Score 47;  DB 8;  Length 516;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3070 GCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        418 GCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 464


RESULT 126
US-09-925-065A-440136/c
; Sequence 440136, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 440136
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-440136

  Query Match             1.5%;  Score 47;  DB 4;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        161 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 115


RESULT 127
US-09-925-065A-440137/c
; Sequence 440137, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 440137
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-440137

  Query Match             1.5%;  Score 47;  DB 4;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        161 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 115


RESULT 128
US-09-925-065A-623360/c
; Sequence 623360, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 623360
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-623360

  Query Match             1.5%;  Score 47;  DB 4;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        336 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 290


RESULT 129
US-10-027-632-115306/c
; Sequence 115306, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115306
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115306

  Query Match             1.5%;  Score 47;  DB 5;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794

.
RESULT 130
US-10-027-632-115307/c
; Sequence 115307, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115307
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115307

  Query Match             1.5%;  Score 47;  DB 5;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794


RESULT 131
US-10-027-632-115308/c
; Sequence 115308, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115308
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115308

  Query Match             1.5%;  Score 47;  DB 5;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794


RESULT 132
US-10-027-632-115306/c
; Sequence 115306, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
```

```
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115306
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115306

  Query Match             1.5%;  Score 47;  DB 6;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794


RESULT 133
US-10-027-632-115307/c
; Sequence 115307, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115307
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115307

  Query Match             1.5%;  Score 47;  DB 6;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794


RESULT 134
US-10-027-632-115308/c
; Sequence 115308, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115308
;   LENGTH: 3231
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115308

  Query Match             1.5%;  Score 47;  DB 6;  Length 3231;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 794


RESULT 135
US-10-719-993-51173/c
; Sequence 51173, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51173
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-51173

  Query Match             1.5%;  Score 46;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 152


RESULT 136
US-10-674-124A-25317/c
; Sequence 25317, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
```

```
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 25317
;   LENGTH: 322
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC007064.27_66627
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 22
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 14352146
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 34953
US-10-674-124A-25317

  Query Match             1.5%;  Score 46;  DB 8;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||
Db        153 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 108


RESULT 137
US-10-914-037-282/c
; Sequence 282, Application US/10914037
; Publication No. US20050003444A1
; GENERAL INFORMATION:
;  APPLICANT: Nehls, Michael
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN POLYNUCLEOTIDES AND THE
;  TITLE OF INVENTION:  POLYPEPTIDES ENCODED THEREBY
;  FILE REFERENCE: 8535-0029-999
;  CURRENT APPLICATION NUMBER: US/10/914,037
;  CURRENT FILING DATE:  2004-08-06
;  PRIOR APPLICATION NUMBER: US/09/428,674
;  PRIOR FILING DATE: 1999-10-27
;  PRIOR APPLICATION NUMBER: US 60/106,442
;  PRIOR FILING DATE: 1998-10-30
;  NUMBER OF SEQ ID NOS: 1008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 282
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(380)
;   OTHER INFORMATION: n = A,T,C or G
US-10-914-037-282

  Query Match             1.5%;  Score 46;  DB 8;  Length 380;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db         49 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 4


RESULT 138
US-09-867-701-6376/c
; Sequence 6376, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6376
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-6376

  Query Match             1.5%;  Score 46;  DB 3;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 61


RESULT 139
US-10-674-124A-24232/c
; Sequence 24232, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 24232
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC020909.4_102225
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 19
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 67910743
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
```

```
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 99683
US-10-674-124A-24232

  Query Match             1.5%;  Score 46;  DB 8;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        335 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 290


RESULT 140
US-09-814-353-17260/c
; Sequence 17260, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17260
;   LENGTH: 458
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-814-353-17260

  Query Match             1.5%;  Score 46;  DB 3;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        405 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 360


RESULT 141
US-10-674-124A-16717
; Sequence 16717, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 16717
;   LENGTH: 482
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC016744.5_76819
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 10
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 111450655
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 136958
US-10-674-124A-16717

  Query Match             1.5%;  Score 46;  DB 8;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 357


RESULT 142
US-10-357-930-60554
; Sequence 60554, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60554
;   LENGTH: 497
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-60554

  Query Match             1.5%;  Score 46;  DB 8;  Length 497;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 237


RESULT 143
US-10-027-632-91304
; Sequence 91304, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91304
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91304

  Query Match             1.5%;  Score 46;  DB 5;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 144
US-10-027-632-91305
; Sequence 91305, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91305
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91305

  Query Match             1.5%;  Score 46;  DB 5;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 145
US-10-027-632-91306
; Sequence 91306, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91306
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91306

  Query Match             1.5%;  Score 46;  DB 5;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 146
US-10-027-632-91304
; Sequence 91304, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91304
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91304

  Query Match             1.5%;  Score 46;  DB 6;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 147
US-10-027-632-91305
; Sequence 91305, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91305
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91305

  Query Match             1.5%;  Score 46;  DB 6;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 148
US-10-027-632-91306
; Sequence 91306, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 91306
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-91306

  Query Match             1.5%;  Score 46;  DB 6;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 428


RESULT 149
US-09-925-065A-124249
; Sequence 124249, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 124249
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-124249

  Query Match             1.5%;  Score 46;  DB 4;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        358 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 403

RESULT 150
US-09-925-065A-843811/c
; Sequence 843811, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 843811
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-843811

  Query Match             1.5%;  Score 46;  DB 4;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        298 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 253


RESULT 151
US-09-925-065A-801712/c
; Sequence 801712, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 801712
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-801712

  Query Match             1.5%;  Score 46;  DB 4;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 254


RESULT 152
US-09-925-065A-48663/c
; Sequence 48663, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48663
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-48663

  Query Match             1.5%;  Score 46;  DB 4;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 495


RESULT 153
US-09-925-065A-65044
; Sequence 65044, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65044
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-65044
```

```
  Query Match             1.5%;  Score 46;  DB 4;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 285


RESULT 154
US-09-925-065A-47410/c
; Sequence 47410, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47410
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-47410

  Query Match             1.5%;  Score 46;  DB 4;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 155
US-09-925-065A-47411/c
; Sequence 47411, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47411
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-47411

  Query Match             1.5%;  Score 46;  DB 4;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 156
US-09-925-065A-930543/c
; Sequence 930543, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 930543
;   LENGTH: 615
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-930543

  Query Match             1.5%;  Score 46;  DB 4;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        290 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 245


RESULT 157
US-09-925-065A-930544/c
; Sequence 930544, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 930544
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930544

Query Match 1.5%; Score 46; DB 4; Length 615;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||||
Db 290 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 245

RESULT 158
US-09-925-065A-870312/c
; Sequence 870312, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870312
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-870312

Query Match 1.5%; Score 46; DB 4; Length 616;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||||
Db 246 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 201

RESULT 159
US-10-027-632-115128
; Sequence 115128, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115128
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115128

Query Match 1.5%; Score 46; DB 5; Length 621;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||||
Db 364 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 409

RESULT 160
US-10-027-632-115129
; Sequence 115129, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115129
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115129

Query Match 1.5%; Score 46; DB 5; Length 621;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||||
Db 364 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 409

RESULT 161
US-10-027-632-115128
; Sequence 115128, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

```
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115128
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115128

  Query Match             1.5%;  Score 46;  DB 6;  Length 621;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 409


RESULT 162
US-10-027-632-115129
; Sequence 115129, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 115129
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-115129

  Query Match             1.5%;  Score 46;  DB 6;  Length 621;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 409


RESULT 163
US-09-925-065A-554324
; Sequence 554324, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 554324
;   LENGTH: 654
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-554324

  Query Match             1.5%;  Score 46;  DB 4;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 199


RESULT 164
US-09-925-065A-759345/c
; Sequence 759345, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 759345
;   LENGTH: 654
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-759345

  Query Match             1.5%;  Score 46;  DB 4;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 229


RESULT 165
US-09-925-065A-790726
; Sequence 790726, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790726
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-790726

  Query Match             1.5%;  Score 46;  DB 4;  Length 659;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 225


RESULT 166
US-09-925-065A-790727
; Sequence 790727, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790727
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-790727

  Query Match             1.5%;  Score 46;  DB 4;  Length 659;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 225


RESULT 167
US-09-925-065A-771454
; Sequence 771454, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 771454
;   LENGTH: 661
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-771454

  Query Match             1.5%;  Score 46;  DB 4;  Length 661;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 433


RESULT 168
US-09-925-065A-737520
; Sequence 737520, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737520
;   LENGTH: 672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-09-925-065A-737520

  Query Match             1.5%;  Score 46;  DB 4;  Length 672;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 430


RESULT 169
US-09-925-065A-737227/c
; Sequence 737227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737227
;   LENGTH: 680
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-737227

  Query Match             1.5%;  Score 46;  DB 4;  Length 680;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 170
US-09-925-065A-852555
; Sequence 852555, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 852555
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-852555

  Query Match             1.5%;  Score 46;  DB 4;  Length 689;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 338


RESULT 171
US-09-925-065A-852556
; Sequence 852556, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 852556
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-852556

  Query Match             1.5%;  Score 46;  DB 4;  Length 689;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 338


RESULT 172
US-10-027-632-144717/c
; Sequence 144717, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
```

```
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144717
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144717

  Query Match             1.5%;  Score 46;  DB 5;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 173
US-10-027-632-144718/c
; Sequence 144718, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144718
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144718

  Query Match             1.5%;  Score 46;  DB 5;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 174
US-10-027-632-144719/c
; Sequence 144719, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144719
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144719

  Query Match             1.5%;  Score 46;  DB 5;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 175
US-10-027-632-144717/c
; Sequence 144717, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144717
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144717

  Query Match             1.5%;  Score 46;  DB 6;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249
```

```
RESULT 176
US-10-027-632-144718/c
; Sequence 144718, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144718
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144718

  Query Match             1.5%;  Score 46;  DB 6;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 177
US-10-027-632-144719/c
; Sequence 144719, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144719
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-144719

  Query Match             1.5%;  Score 46;  DB 6;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 178
US-09-925-065A-938958/c
; Sequence 938958, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 938958
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938958

  Query Match             1.5%;  Score 46;  DB 4;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 179
US-09-925-065A-938959/c
; Sequence 938959, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 938959
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938959

  Query Match             1.5%;  Score 46;  DB 4;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 180
US-09-925-065A-938960/c
; Sequence 938960, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 938960
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938960

  Query Match             1.5%;  Score 46;  DB 4;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 181
US-09-925-065A-711495
; Sequence 711495, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 711495
;   LENGTH: 1105
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-711495

  Query Match             1.5%;  Score 46;  DB 4;  Length 1105;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 919


RESULT 182
US-09-925-065A-678054
; Sequence 678054, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 678054
;   LENGTH: 1364
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-678054

  Query Match             1.5%;  Score 46;  DB 4;  Length 1364;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 396


RESULT 183
US-09-925-065A-689300
; Sequence 689300, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 689300
;   LENGTH: 1635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-689300

  Query Match             1.5%;  Score 46;  DB 4;  Length 1635;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        831 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 876


RESULT 184
US-09-925-065A-677066
; Sequence 677066, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 677066
;   LENGTH: 2227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-677066

  Query Match             1.5%;  Score 46;  DB 4;  Length 2227;
  Best Local Similarity   100.0%;  Pred. No. 7e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db       2035 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2080


RESULT 185
US-10-027-632-260094
; Sequence 260094, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 260094
;   LENGTH: 2798
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-260094

  Query Match             1.5%;  Score 46;  DB 5;  Length 2798;
  Best Local Similarity   100.0%;  Pred. No. 7e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 1479


RESULT 186
US-10-027-632-260094
; Sequence 260094, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 260094
;   LENGTH: 2798
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-260094

  Query Match             1.5%;  Score 46;  DB 6;  Length 2798;
  Best Local Similarity   100.0%;  Pred. No. 7e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 1479


RESULT 187
US-10-008-789-10/c
; Sequence 10, Application US/10008789
```

; Publication No. US20030125276A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Kenneth Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF THYROID HORMONE RECEPTOR INTERACTOR 6 EXP
; FILE REFERENCE: RTS-0333
; CURRENT APPLICATION NUMBER: US/10/008,789
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 7001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (486)...(740)
; OTHER INFORMATION: exon 1
; NAME/KEY: exon:intron junction
; LOCATION: (740)...(741)
; OTHER INFORMATION: exon 1:intron 1
; NAME/KEY: intron
; LOCATION: (741)...(994)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (994)...(995)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (995)...(1122)
; OTHER INFORMATION: exon 2
; NAME/KEY: exon:intron junction
; LOCATION: (1122)...(1123)
; OTHER INFORMATION: exon 2:intron 2
; NAME/KEY: intron
; LOCATION: (1123)...(1241)
; OTHER INFORMATION: intron 2
; NAME/KEY: intron:exon junction
; LOCATION: (1241)...(1242)
; OTHER INFORMATION: intron 2:exon 3
; NAME/KEY: exon
; LOCATION: (1242)...(1367)
; OTHER INFORMATION: exon 3
; NAME/KEY: exon:intron junction
; LOCATION: (1367)...(1368)
; OTHER INFORMATION: exon 3:intron 3
; NAME/KEY: intron
; LOCATION: (1368)...(1628)
; OTHER INFORMATION: intron 3
; NAME/KEY: intron:exon junction
; LOCATION: (1628)...(1629)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (1629)...(2000)
; OTHER INFORMATION: exon 4
; NAME/KEY: exon:intron junction
; LOCATION: (2000)...(2001)
; OTHER INFORMATION: exon 4:intron 4
; NAME/KEY: intron
; LOCATION: (2001)...(3503)
; OTHER INFORMATION: intron 4
; NAME/KEY: intron:exon junction
; LOCATION: (3503)...(3504)
; OTHER INFORMATION: intron 4:exon 5
; NAME/KEY: exon
; LOCATION: (3504)...(3597)
; OTHER INFORMATION: exon 5
; NAME/KEY: exon:intron junction
; LOCATION: (3597)...(3598)
; OTHER INFORMATION: exon 5:intron 5
; NAME/KEY: intron
; LOCATION: (3598)...(3707)
; OTHER INFORMATION: intron 5
; NAME/KEY: intron:exon junction
; LOCATION: (3707)...(3708)
; OTHER INFORMATION: intron 5:exon 6
; NAME/KEY: exon
; LOCATION: (3708)...(3877)
; OTHER INFORMATION: exon 6
; NAME/KEY: exon:intron junction
; LOCATION: (3877)...(3878)
; OTHER INFORMATION: exon 6:intron 6
; NAME/KEY: intron
; LOCATION: (3878)...(4676)
; OTHER INFORMATION: intron 6
; NAME/KEY: intron:exon junction
; LOCATION: (4676)...(4677)
; OTHER INFORMATION: intron 6:exon 7
; NAME/KEY: exon
; LOCATION: (4677)...(4855)
; OTHER INFORMATION: exon 7
; NAME/KEY: exon:intron junction
; LOCATION: (4855)...(4856)
; OTHER INFORMATION: exon 7:intron 7
; NAME/KEY: intron
; LOCATION: (4856)...(5757)
; OTHER INFORMATION: intron 7
; NAME/KEY: intron:exon junction
; LOCATION: (5757)...(5758)
; OTHER INFORMATION: intron 7:exon 8
; NAME/KEY: exon
; LOCATION: (5758)...(5878)
; OTHER INFORMATION: exon 8
; NAME/KEY: exon:intron junction
; LOCATION: (5878)...(5879)
; OTHER INFORMATION: exon 8:intron 8
; NAME/KEY: intron
; LOCATION: (5879)...(6305)
; OTHER INFORMATION: intron 8
; NAME/KEY: intron:exon junction
; LOCATION: (6305)...(6306)
; OTHER INFORMATION: intron 8:exon 9
; NAME/KEY: exon
; LOCATION: (6306)...(6650)
; OTHER INFORMATION: exon 9
US-10-008-789-10

Query Match 1.5%; Score 46; DB 6; Length 7001;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||||
Db 6054 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 6009

RESULT 188
US-11-071-724-10/c
; Sequence 10, Application US/11071724
; Publication No. US20050222073A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Brett P. Monia
; APPLICANT: Rosanne Crooke
; APPLICANT: Mark Graham
; APPLICANT: Susan M. Freier
; APPLICANT: Pamela Nero
; APPLICANT: Edward Wancewicz
; APPLICANT: Nicholas M. Dean
; APPLICANT: Lex M. Cowsert
; APPLICANT: Kenneth Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF THYROID HORMONE RECEPTOR INTERACTOR 6 EXP
; FILE REFERENCE: BNDL-0018US.P1
; CURRENT APPLICATION NUMBER: US/11/071,724
; CURRENT FILING DATE: 2005-03-02
; PRIOR APPLICATION NUMBER: 10/190,366
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 10/008,789

; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 11/031,827
; PRIOR FILING DATE: 2005-01-07
; PRIOR APPLICATION NUMBER: 09/925,139
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 10/114,544
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 10/019,470
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/19019
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 09/357,071
; PRIOR FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 7001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (486)...(740)
; OTHER INFORMATION: exon 1
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (740)...(741)
; OTHER INFORMATION: exon 1:intron 1
; FEATURE:
; NAME/KEY: intron
; LOCATION: (741)...(994)
; OTHER INFORMATION: intron 1
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (994)...(995)
; OTHER INFORMATION: intron 1:exon 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: (995)...(1122)
; OTHER INFORMATION: exon 2
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (1122)...(1123)
; OTHER INFORMATION: exon 2:intron 2
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1123)...(1241)
; OTHER INFORMATION: intron 2
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (1241)...(1242)
; OTHER INFORMATION: intron 2:exon 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1242)...(1367)
; OTHER INFORMATION: exon 3
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (1367)...(1368)
; OTHER INFORMATION: exon 3:intron 3
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1368)...(1628)
; OTHER INFORMATION: intron 3
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (1628)...(1629)
; OTHER INFORMATION: intron 3:exon 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1629)...(2000)
; OTHER INFORMATION: exon 4
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (2000)...(2001)
; OTHER INFORMATION: exon 4:intron 4
; FEATURE:
; NAME/KEY: intron
; LOCATION: (2001)...(3503)
; OTHER INFORMATION: intron 4
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (3503)...(3504)
; OTHER INFORMATION: intron 4:exon 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3504)...(3597)
; OTHER INFORMATION: exon 5
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (3597)...(3598)
; OTHER INFORMATION: exon 5:intron 5
; FEATURE:
; NAME/KEY: intron
; LOCATION: (3598)...(3707)
; OTHER INFORMATION: intron 5
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (3707)...(3708)
; OTHER INFORMATION: intron 5:exon 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3708)...(3877)
; OTHER INFORMATION: exon 6
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (3877)...(3878)
; OTHER INFORMATION: exon 6:intron 6
; FEATURE:
; NAME/KEY: intron
; LOCATION: (3878)...(4676)
; OTHER INFORMATION: intron 6
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (4676)...(4677)
; OTHER INFORMATION: intron 6:exon 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: (4677)...(4855)
; OTHER INFORMATION: exon 7
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (4855)...(4856)
; OTHER INFORMATION: exon 7:intron 7
; FEATURE:
; NAME/KEY: intron
; LOCATION: (4856)...(5757)
; OTHER INFORMATION: intron 7
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (5757)...(5758)
; OTHER INFORMATION: intron 7:exon 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: (5758)...(5878)
; OTHER INFORMATION: exon 8
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (5878)...(5879)
; OTHER INFORMATION: exon 8:intron 8
; FEATURE:
; NAME/KEY: intron
; LOCATION: (5879)...(6305)
; OTHER INFORMATION: intron 8
; FEATURE:
; NAME/KEY: intron:exon junction
; LOCATION: (6305)...(6306)
; OTHER INFORMATION: intron 8:exon 9

```
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (6306)...(6650)
;   OTHER INFORMATION: exon 9
US-11-071-724-10

  Query Match             1.5%;  Score 46;  DB 10;  Length 7001;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       6054 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 6009


RESULT 189
US-09-764-891-9601
; Sequence 9601, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9601
;   LENGTH: 13409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-9601

  Query Match             1.5%;  Score 46;  DB 3;  Length 13409;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      12121 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 12166


RESULT 190
US-10-205-428-930
; Sequence 930, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA117C1
;  CURRENT APPLICATION NUMBER: US/10/205,428
;  CURRENT FILING DATE:  2002-07-26
;  PRIOR APPLICATION NUMBER: 09/764,892
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1019
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 930
;   LENGTH: 13409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-205-428-930

  Query Match             1.5%;  Score 46;  DB 5;  Length 13409;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      12121 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 12166


RESULT 191
US-09-764-847-1916
; Sequence 1916, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC009
;  CURRENT APPLICATION NUMBER: US/09/764,847
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 2003
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1916
;   LENGTH: 18501
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (9022)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (9023)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-847-1916

  Query Match             1.5%;  Score 46;  DB 3;  Length 18501;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 269


RESULT 192
US-10-092-154-1916
; Sequence 1916, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC009C1
;  CURRENT APPLICATION NUMBER: US/10/092,154
;  CURRENT FILING DATE:  2002-03-07
;  NUMBER OF SEQ ID NOS: 2003
;  Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1916
;   LENGTH: 18501
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9022)
```

```
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (9023)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-092-154-1916

  Query Match             1.5%;  Score 46;  DB 5;  Length 18501;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 269


RESULT 193
US-10-187-659A-11/c
; Sequence 11, Application US/10187659A
; Publication No. US20040002152A1
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  TITLE OF INVENTION: ANTISENSE MODULATION OF P2X4 EXPRESSION
;  FILE REFERENCE: RTS-0379
;  CURRENT APPLICATION NUMBER: US/10/187,659A
;  CURRENT FILING DATE:  2002-07-01
;  NUMBER OF SEQ ID NOS: 143
; SEQ ID NO 11
;   LENGTH: 25001
;   TYPE: DNA
;   ORGANISM: H. sapiens
US-10-187-659A-11

  Query Match             1.5%;  Score 46;  DB 6;  Length 25001;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2588


RESULT 194
US-10-455-552-1/c
; Sequence 1, Application US/10455552
; Publication No. US20040018533A1
; GENERAL INFORMATION:
;  APPLICANT: Adam, Gail Isabel
;  APPLICANT:  Langdown, Maria
;  APPLICANT:  Roth, Richard
;  APPLICANT:  Denissenko, Mikhail
;  APPLICANT:  Smylie, Kevin
;  TITLE OF INVENTION: DIAGNOSING PREDISPOSITION TO FAT
;  TITLE OF INVENTION:  DEPOSITION AND THERAPEUTIC METHODS FOR REDUCING FAT
;  TITLE OF INVENTION:  DEPOSITION AND TREATMENT OF ASSOCIATED CONDITIONS
;  FILE REFERENCE: 52459-20030.00
;  CURRENT APPLICATION NUMBER: US/10/455,552
;  CURRENT FILING DATE:  2003-06-04
;  PRIOR APPLICATION NUMBER: US 60/386,012
;  PRIOR FILING DATE: 2002-06-04
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 28616
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-455-552-1

  Query Match             1.5%;  Score 46;  DB 7;  Length 28616;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       4783 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 4738


RESULT 195
US-10-719-993-7030/c
; Sequence 7030, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7030
;   LENGTH: 57095
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-7030

  Query Match             1.5%;  Score 46;  DB 8;  Length 57095;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 23227


RESULT 196
US-10-322-281-526/c
; Sequence 526, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001000
;  CURRENT APPLICATION NUMBER: US/10/322,281
;  CURRENT FILING DATE:  2002-12-17
;  NUMBER OF SEQ ID NOS: 866
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 526
;   LENGTH: 58922
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(58922)
;   OTHER INFORMATION: n = A,T,C or G
US-10-322-281-526

  Query Match             1.5%;  Score 46;  DB 7;  Length 58922;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      53826 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 53781


RESULT 197
US-10-741-601-5619/c
; Sequence 5619, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
```

```
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5619
;    LENGTH: 59914
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(59914)
;    OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-601-5619

  Query Match             1.5%;  Score 46;  DB 7;  Length 59914;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      10387 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 10342


RESULT 198
US-10-719-993-6871/c
; Sequence 6871, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6871
;    LENGTH: 93011
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(93011)
;    OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6871

  Query Match             1.5%;  Score 46;  DB 8;  Length 93011;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      82157 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 82112


RESULT 199
US-10-052-482-232
; Sequence 232, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
;  APPLICANT: Engelhard, Eric
;  APPLICANT:  Morris, David
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71087/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/10/052,482
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 232
;    LENGTH: 96595
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (49183)..(49621)
;    OTHER INFORMATION: "n" at positions 49183 to 49621 can be any base
US-10-052-482-232

  Query Match             1.5%;  Score 46;  DB 7;  Length 96595;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      57090 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 57135


RESULT 200
US-10-704-513-1/c
; Sequence 1, Application US/10704513
; Publication No. US20050170500A1
; GENERAL INFORMATION:
;  APPLICANT: ROTH, RICHARD B.
;  APPLICANT:  NELSON, MATTHEW ROBERTS
;  APPLICANT:  KAMMERER, STEFAN M.
;  APPLICANT:  BRAUN, ANDREAS
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF MELANOMA AND TREATMENTS
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: SEQ-4062-UT
;  CURRENT APPLICATION NUMBER: US/10/704,513
;  CURRENT FILING DATE:  2003-11-06
;  PRIOR APPLICATION NUMBER: 60/489,703
;  PRIOR FILING DATE: 2003-07-23
;  PRIOR APPLICATION NUMBER: 60/424,475
;  PRIOR FILING DATE: 2002-11-06
;  NUMBER OF SEQ ID NOS: 774
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;    LENGTH: 99100
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: modified_base
;    LOCATION: (47431)
;    OTHER INFORMATION: a, c, g, t, unknown, or other
US-10-704-513-1

  Query Match             1.5%;  Score 46;  DB 9;  Length 99100;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      16521 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 16476


Search completed: May 11, 2006, 06:40:25
Job time : 2402 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:54:19 ; Search time 1343 Seconds
(without alignments)
9481.115 Million cell updates/sec

Title: US-10-760-320A-102
Perfect score: 3122
Sequence: 1 actagaggtggggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9312410 seqs, 2039259788 residues

Word size : 1

Total number of hits satisfying chosen parameters: 18623998

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq1:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
8: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
11: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
12: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
13: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq4:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
16: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
17: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
18: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq5:*
19: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	509	16.3	560	7	US-09-925-065A-740956	Sequence 740956,
c	2	506	16.2	557	7	US-09-925-065A-769508	Sequence 769508,
c	3	505	16.2	556	7	US-09-925-065A-736351	Sequence 736351,
	4	504	16.1	555	7	US-09-925-065A-737120	Sequence 737120,
c	5	53	1.7	477	11	US-10-301-480-20199	Sequence 20199, A
c	6	53	1.7	477	12	US-10-301-480-633608	Sequence 633608,
c	7	53	1.7	492	11	US-10-301-480-20198	Sequence 20198, A
c	8	53	1.7	492	12	US-10-301-480-633607	Sequence 633607,
c	9	52	1.7	559	7	US-09-925-065A-591947	Sequence 591947,
	10	52	1.7	574	7	US-09-925-065A-768118	Sequence 768118,
c	11	52	1.7	592	7	US-09-925-065A-59682	Sequence 59682, A
c	12	52	1.7	592	11	US-10-301-480-160920	Sequence 160920,
c	13	52	1.7	592	12	US-10-301-480-774329	Sequence 774329,
c	14	52	1.7	614	12	US-10-301-480-454942	Sequence 454942,
c	15	52	1.7	614	12	US-10-301-480-1068351	Sequence 1068351,
c	16	52	1.7	617	7	US-09-925-065A-386361	Sequence 386361,
c	17	52	1.7	618	12	US-10-301-480-611724	Sequence 611724,
c	18	52	1.7	618	12	US-10-301-480-1225133	Sequence 1225133,
c	19	52	1.7	695	7	US-09-925-065A-708701	Sequence 708701,
	20	52	1.7	974	12	US-10-301-480-598972	Sequence 598972,
	21	52	1.7	974	12	US-10-301-480-1212381	Sequence 1212381,
	22	52	1.7	118996	17	US-11-121-086-84	Sequence 84, Appl
	23	52	1.7	137671	17	US-11-121-086-47	Sequence 47, Appl
c	24	52	1.7	398287	10	US-10-995-561-13396	Sequence 13396, A
c	25	51	1.6	385	7	US-09-925-065A-211249	Sequence 211249,
	26	51	1.6	561	7	US-09-925-065A-141833	Sequence 141833,
	27	51	1.6	561	12	US-10-301-480-236672	Sequence 236672,
	28	51	1.6	561	12	US-10-301-480-850081	Sequence 850081,
	29	51	1.6	587	11	US-10-301-480-1011	Sequence 1011, Ap
	30	51	1.6	587	11	US-10-301-480-42180	Sequence 42180, A
	31	51	1.6	587	12	US-10-301-480-614420	Sequence 614420,
	32	51	1.6	587	12	US-10-301-480-655589	Sequence 655589,
c	33	51	1.6	852	12	US-10-301-480-589593	Sequence 589593,
c	34	51	1.6	852	12	US-10-301-480-1203002	Sequence 1203002,
c	35	51	1.6	197096	17	US-11-121-086-107	Sequence 107, App
c	36	50	1.6	571	7	US-09-925-065A-484883	Sequence 484883,
c	37	50	1.6	572	7	US-09-925-065A-785967	Sequence 785967,
c	38	50	1.6	625	7	US-09-925-065A-894777	Sequence 894777,
c	39	50	1.6	630	7	US-09-925-065A-892664	Sequence 892664,
c	40	50	1.6	633	7	US-09-925-065A-892405	Sequence 892405,
c	41	50	1.6	633	7	US-09-925-065A-892514	Sequence 892514,
	42	50	1.6	633	7	US-09-925-065A-949927	Sequence 949927,
	43	50	1.6	676	7	US-09-925-065A-917225	Sequence 917225,
	44	50	1.6	676	7	US-09-925-065A-917226	Sequence 917226,
	45	50	1.6	676	7	US-09-925-065A-944569	Sequence 944569,
	46	50	1.6	739	7	US-09-925-065A-934608	Sequence 934608,
	47	50	1.6	742	7	US-09-925-065A-941367	Sequence 941367,
	48	50	1.6	999	12	US-10-301-480-567391	Sequence 567391,
	49	50	1.6	999	12	US-10-301-480-1180800	Sequence 1180800,
	50	50	1.6	1986	11	US-10-301-480-100096	Sequence 100096,
	51	50	1.6	1986	12	US-10-301-480-713505	Sequence 713505,
c	52	50	1.6	187745	17	US-11-121-086-83	Sequence 83, Appl
	53	50	1.6	207908	17	US-11-112-908-21	Sequence 21, Appl
	54	50	1.6	212805	17	US-11-112-908-19	Sequence 19, Appl
c	55	49	1.6	417	12	US-10-301-480-235645	Sequence 235645,
c	56	49	1.6	417	12	US-10-301-480-849054	Sequence 849054,
c	57	49	1.6	420	7	US-09-925-065A-140670	Sequence 140670,
c	58	49	1.6	530	11	US-10-301-480-15264	Sequence 15264, A
c	59	49	1.6	530	12	US-10-301-480-235646	Sequence 235646,
c	60	49	1.6	530	12	US-10-301-480-628673	Sequence 628673,
c	61	49	1.6	530	12	US-10-301-480-849055	Sequence 849055,
c	62	49	1.6	532	7	US-09-925-065A-140671	Sequence 140671,
c	63	49	1.6	552	7	US-09-925-065A-18609	Sequence 18609, A
c	64	49	1.6	552	7	US-09-925-065A-18610	Sequence 18610, A
c	65	49	1.6	552	7	US-09-925-065A-18611	Sequence 18611, A
c	66	49	1.6	552	11	US-10-301-480-119846	Sequence 119846,
c	67	49	1.6	552	11	US-10-301-480-119847	Sequence 119847,
c	68	49	1.6	552	11	US-10-301-480-119848	Sequence 119848,
c	69	49	1.6	552	12	US-10-301-480-733255	Sequence 733255,
c	70	49	1.6	552	12	US-10-301-480-733256	Sequence 733256,
c	71	49	1.6	552	12	US-10-301-480-733257	Sequence 733257,
c	72	49	1.6	695	7	US-09-925-065A-708700	Sequence 708700,
c	73	49	1.6	750	7	US-09-925-065A-932035	Sequence 932035,
c	74	49	1.6	750	7	US-09-925-065A-932036	Sequence 932036,
c	75	49	1.6	243428	13	US-10-508-307-14	Sequence 14, Appl
	76	48	1.5	437	11	US-10-301-480-24305	Sequence 24305, A
	77	48	1.5	437	12	US-10-301-480-637714	Sequence 637714,
	78	48	1.5	438	11	US-10-301-480-43148	Sequence 43148, A
	79	48	1.5	438	12	US-10-301-480-656557	Sequence 656557,
c	80	48	1.5	617	12	US-10-301-480-597620	Sequence 597620,
c	81	48	1.5	617	12	US-10-301-480-1211029	Sequence 1211029,
c	82	48	1.5	802	12	US-10-301-480-597621	Sequence 597621,
c	83	48	1.5	802	12	US-10-301-480-1211030	Sequence 1211030,
	84	48	1.5	19675	17	US-11-124-368A-2878	Sequence 2878, Ap
c	85	47	1.5	534	7	US-09-925-065A-440136	Sequence 440136,
c	86	47	1.5	534	7	US-09-925-065A-440137	Sequence 440137,
c	87	47	1.5	566	7	US-09-925-065A-623360	Sequence 623360,

	88	46	1.5	201	10	US-10-995-561-64399	Sequence 64399, A
c	89	46	1.5	496	11	US-10-301-480-20200	Sequence 20200, A
c	90	46	1.5	496	12	US-10-301-480-633609	Sequence 633609,
c	91	46	1.5	514	12	US-10-301-480-570118	Sequence 570118,
c	92	46	1.5	514	12	US-10-301-480-1183527	Sequence 1183527,
c	93	46	1.5	540	12	US-10-301-480-577517	Sequence 577517,
c	94	46	1.5	540	12	US-10-301-480-1190926	Sequence 1190926,
	95	46	1.5	559	7	US-09-925-065A-124249	Sequence 124249,
c	96	46	1.5	568	12	US-10-301-480-221987	Sequence 221987,
c	97	46	1.5	568	12	US-10-301-480-835396	Sequence 835396,
c	98	46	1.5	606	7	US-09-925-065A-843811	Sequence 843811,
c	99	46	1.5	609	7	US-09-925-065A-801712	Sequence 801712,
c	100	46	1.5	610	7	US-09-925-065A-48663	Sequence 48663, A
	101	46	1.5	610	7	US-09-925-065A-65044	Sequence 65044, A
c	102	46	1.5	610	11	US-10-301-480-149901	Sequence 149901,
	103	46	1.5	610	11	US-10-301-480-166282	Sequence 166282,
c	104	46	1.5	610	12	US-10-301-480-763310	Sequence 763310,
	105	46	1.5	610	12	US-10-301-480-779691	Sequence 779691,
c	106	46	1.5	614	7	US-09-925-065A-47410	Sequence 47410, A
c	107	46	1.5	614	7	US-09-925-065A-47411	Sequence 47411, A
c	108	46	1.5	614	11	US-10-301-480-148648	Sequence 148648,
c	109	46	1.5	614	11	US-10-301-480-148649	Sequence 148649,
c	110	46	1.5	614	12	US-10-301-480-762057	Sequence 762057,
c	111	46	1.5	614	12	US-10-301-480-762058	Sequence 762058,
c	112	46	1.5	615	7	US-09-925-065A-930543	Sequence 930543,
c	113	46	1.5	615	7	US-09-925-065A-930544	Sequence 930544,
c	114	46	1.5	616	7	US-09-925-065A-870312	Sequence 870312,
	115	46	1.5	654	7	US-09-925-065A-554324	Sequence 554324,
c	116	46	1.5	654	7	US-09-925-065A-759345	Sequence 759345,
	117	46	1.5	659	7	US-09-925-065A-790726	Sequence 790726,
	118	46	1.5	659	7	US-09-925-065A-790727	Sequence 790727,
	119	46	1.5	661	7	US-09-925-065A-771454	Sequence 771454,
	120	46	1.5	672	7	US-09-925-065A-737520	Sequence 737520,
c	121	46	1.5	680	7	US-09-925-065A-737227	Sequence 737227,
	122	46	1.5	689	7	US-09-925-065A-852555	Sequence 852555,
	123	46	1.5	689	7	US-09-925-065A-852556	Sequence 852556,
c	124	46	1.5	840	7	US-09-925-065A-938958	Sequence 938958,
c	125	46	1.5	840	7	US-09-925-065A-938959	Sequence 938959,
c	126	46	1.5	840	7	US-09-925-065A-938960	Sequence 938960,
c	127	46	1.5	989	12	US-10-301-480-601877	Sequence 601877,
c	128	46	1.5	989	12	US-10-301-480-1215286	Sequence 1215286,
c	129	46	1.5	998	12	US-10-301-480-612422	Sequence 612422,
c	130	46	1.5	998	12	US-10-301-480-1225831	Sequence 1225831,
	131	46	1.5	1105	7	US-09-925-065A-711495	Sequence 711495,
	132	46	1.5	1364	7	US-09-925-065A-678054	Sequence 678054,
	133	46	1.5	1635	7	US-09-925-065A-689300	Sequence 689300,
	134	46	1.5	2227	7	US-09-925-065A-677066	Sequence 677066,
	135	46	1.5	2798	11	US-10-301-480-94903	Sequence 94903, A
	136	46	1.5	2798	12	US-10-301-480-708312	Sequence 708312,
	137	46	1.5	56448	10	US-10-995-561-13369	Sequence 13369, A
c	138	46	1.5	98345	17	US-11-112-908-36	Sequence 36, Appl
	139	46	1.5	109661	11	US-10-330-773-795	Sequence 795, App
c	140	46	1.5	150450	17	US-11-112-908-54	Sequence 54, Appl
c	141	46	1.5	191343	17	US-11-112-908-53	Sequence 53, Appl
c	142	45	1.4	201	10	US-10-995-561-58923	Sequence 58923, A
c	143	45	1.4	201	10	US-10-995-561-58959	Sequence 58959, A
	144	45	1.4	201	10	US-10-995-561-68537	Sequence 68537, A
c	145	45	1.4	524	11	US-10-301-480-49738	Sequence 49738, A
c	146	45	1.4	524	11	US-10-301-480-49739	Sequence 49739, A
c	147	45	1.4	524	12	US-10-301-480-663147	Sequence 663147,
c	148	45	1.4	524	12	US-10-301-480-663148	Sequence 663148,
	149	45	1.4	527	11	US-10-301-480-9096	Sequence 9096, Ap
	150	45	1.4	527	12	US-10-301-480-622505	Sequence 622505,
	151	45	1.4	545	7	US-09-925-065A-765896	Sequence 765896,
c	152	45	1.4	546	7	US-09-925-065A-799095	Sequence 799095,
	153	45	1.4	549	7	US-09-925-065A-615308	Sequence 615308,
c	154	45	1.4	559	7	US-09-925-065A-591948	Sequence 591948,
c	155	45	1.4	563	7	US-09-925-065A-12438	Sequence 12438, A
c	156	45	1.4	563	11	US-10-301-480-113675	Sequence 113675,
c	157	45	1.4	563	12	US-10-301-480-727084	Sequence 727084,
	158	45	1.4	568	7	US-09-925-065A-13098	Sequence 13098, A
	159	45	1.4	568	7	US-09-925-065A-338452	Sequence 338452,
	160	45	1.4	568	11	US-10-301-480-114335	Sequence 114335,
	161	45	1.4	568	12	US-10-301-480-411598	Sequence 411598,
	162	45	1.4	568	12	US-10-301-480-727744	Sequence 727744,
	163	45	1.4	568	12	US-10-301-480-1025007	Sequence 1025007,
c	164	45	1.4	590	7	US-09-925-065A-371064	Sequence 371064,
	165	45	1.4	592	7	US-09-925-065A-927764	Sequence 927764,
	166	45	1.4	592	7	US-09-925-065A-927765	Sequence 927765,
	167	45	1.4	592	7	US-09-925-065A-949474	Sequence 949474,
c	168	45	1.4	596	7	US-09-925-065A-695847	Sequence 695847,
c	169	45	1.4	598	7	US-09-925-065A-10892	Sequence 10892, A
c	170	45	1.4	598	11	US-10-301-480-112129	Sequence 112129,
c	171	45	1.4	598	12	US-10-301-480-725538	Sequence 725538,
c	172	45	1.4	603	12	US-10-301-480-253035	Sequence 253035,
c	173	45	1.4	603	12	US-10-301-480-866444	Sequence 866444,
c	174	45	1.4	604	7	US-09-925-065A-160227	Sequence 160227,
c	175	45	1.4	604	7	US-09-925-065A-799096	Sequence 799096,
c	176	45	1.4	605	7	US-09-925-065A-795229	Sequence 795229,
c	177	45	1.4	605	7	US-09-925-065A-850785	Sequence 850785,
c	178	45	1.4	606	12	US-10-301-480-441075	Sequence 441075,
c	179	45	1.4	606	12	US-10-301-480-1054484	Sequence 1054484,
c	180	45	1.4	624	7	US-09-925-065A-846491	Sequence 846491,
c	181	45	1.4	627	7	US-09-925-065A-839499	Sequence 839499,
c	182	45	1.4	662	7	US-09-925-065A-675355	Sequence 675355,
c	183	45	1.4	662	7	US-09-925-065A-675356	Sequence 675356,
	184	45	1.4	713	11	US-10-301-480-34229	Sequence 34229, A
	185	45	1.4	713	12	US-10-301-480-647638	Sequence 647638,
c	186	45	1.4	723	12	US-10-301-480-593973	Sequence 593973,
c	187	45	1.4	723	12	US-10-301-480-1207382	Sequence 1207382,
	188	45	1.4	745	12	US-10-301-480-572562	Sequence 572562,
	189	45	1.4	745	12	US-10-301-480-1185971	Sequence 1185971,
c	190	45	1.4	793	7	US-09-925-065A-13524	Sequence 13524, A
c	191	45	1.4	793	7	US-09-925-065A-77588	Sequence 77588, A
c	192	45	1.4	793	11	US-10-301-480-114761	Sequence 114761,
c	193	45	1.4	793	11	US-10-301-480-178827	Sequence 178827,
c	194	45	1.4	793	12	US-10-301-480-728170	Sequence 728170,
c	195	45	1.4	793	12	US-10-301-480-792236	Sequence 792236,
c	196	45	1.4	803	12	US-10-301-480-556511	Sequence 556511,
c	197	45	1.4	803	12	US-10-301-480-1169920	Sequence 1169920,
c	198	45	1.4	818	12	US-10-301-480-556318	Sequence 556318,
c	199	45	1.4	818	12	US-10-301-480-1169727	Sequence 1169727,
	200	45	1.4	896	7	US-09-925-065A-88887	Sequence 88887, A
	201	45	1.4	896	11	US-10-301-480-190128	Sequence 190128,
	202	45	1.4	896	12	US-10-301-480-803537	Sequence 803537,
	203	45	1.4	981	12	US-10-301-480-582545	Sequence 582545,
	204	45	1.4	981	12	US-10-301-480-1195954	Sequence 1195954,
	205	45	1.4	994	12	US-10-301-480-600766	Sequence 600766,
c	206	45	1.4	994	12	US-10-301-480-600775	Sequence 600775,
	207	45	1.4	994	12	US-10-301-480-1214175	Sequence 1214175,
c	208	45	1.4	994	12	US-10-301-480-1214184	Sequence 1214184,
c	209	45	1.4	997	12	US-10-301-480-535489	Sequence 535489,
c	210	45	1.4	997	12	US-10-301-480-551651	Sequence 551651,
c	211	45	1.4	997	12	US-10-301-480-1148898	Sequence 1148898,
c	212	45	1.4	997	12	US-10-301-480-1165060	Sequence 1165060,
	213	45	1.4	1000	12	US-10-301-480-608404	Sequence 608404,
	214	45	1.4	1000	12	US-10-301-480-1221813	Sequence 1221813,
	215	45	1.4	1027	7	US-09-925-065A-85535	Sequence 85535, A
	216	45	1.4	1027	7	US-09-925-065A-85536	Sequence 85536, A
	217	45	1.4	1027	11	US-10-301-480-186775	Sequence 186775,
	218	45	1.4	1027	11	US-10-301-480-186776	Sequence 186776,
	219	45	1.4	1027	12	US-10-301-480-800184	Sequence 800184,
	220	45	1.4	1027	12	US-10-301-480-800185	Sequence 800185,
c	221	45	1.4	1233	11	US-10-301-480-31318	Sequence 31318, A
c	222	45	1.4	1233	11	US-10-301-480-31319	Sequence 31319, A
c	223	45	1.4	1233	12	US-10-301-480-644727	Sequence 644727,
c	224	45	1.4	1233	12	US-10-301-480-644728	Sequence 644728,
c	225	45	1.4	1302	7	US-09-925-065A-553464	Sequence 553464,
	226	45	1.4	1302	11	US-10-301-480-32413	Sequence 32413, A
	227	45	1.4	1302	11	US-10-301-480-32414	Sequence 32414, A
	228	45	1.4	1302	12	US-10-301-480-645822	Sequence 645822,
	229	45	1.4	1302	12	US-10-301-480-645823	Sequence 645823,
	230	45	1.4	1305	7	US-09-925-065A-34751	Sequence 34751, A
	231	45	1.4	1305	11	US-10-301-480-135989	Sequence 135989,
	232	45	1.4	1305	12	US-10-301-480-749398	Sequence 749398,
	233	45	1.4	1498	7	US-09-925-065A-71613	Sequence 71613, A

	234	45	1.4	1498	11	US-10-301-480-172852 Sequence 172852,
	235	45	1.4	1498	12	US-10-301-480-786261 Sequence 786261,
c	236	45	1.4	1792	7	US-09-925-065A-549910 Sequence 549910,
c	237	45	1.4	1792	12	US-10-301-480-527860 Sequence 527860,
c	238	45	1.4	1792	12	US-10-301-480-1141269 Sequence 1141269,
c	239	45	1.4	1869	7	US-09-925-065A-689963 Sequence 689963,
c	240	45	1.4	28693	10	US-10-995-561-13341 Sequence 13341, A
c	241	45	1.4	98560	10	US-10-995-561-13323 Sequence 13323, A
c	242	45	1.4	160226	17	US-11-121-086-29 Sequence 29, Appl
c	243	45	1.4	246960	17	US-11-121-086-8 Sequence 8, Appli
c	244	45	1.4	285300	10	US-10-857-780-6 Sequence 6, Appli
c	245	45	1.4	1080000	10	US-10-928-446A-1 Sequence 1, Appli
c	246	45	1.4	1080000	10	US-10-928-446A-181 Sequence 181, App
c	247	45	1.4	1080000	10	US-10-928-446A-183 Sequence 183, App
c	248	45	1.4	1080000	10	US-10-928-446A-185 Sequence 185, App
c	249	45	1.4	1080000	10	US-10-928-446A-187 Sequence 187, App
c	250	45	1.4	1080000	10	US-10-928-446A-189 Sequence 189, App
c	251	45	1.4	1080000	10	US-10-928-446A-191 Sequence 191, App
c	252	45	1.4	1080000	10	US-10-928-446A-193 Sequence 193, App
c	253	45	1.4	1080000	10	US-10-928-446A-195 Sequence 195, App
c	254	45	1.4	1080000	10	US-10-928-446A-197 Sequence 197, App
c	255	45	1.4	1080000	10	US-10-928-446A-199 Sequence 199, App
c	256	45	1.4	1080000	10	US-10-928-446A-201 Sequence 201, App
c	257	45	1.4	1691140	17	US-11-091-018-1 Sequence 1, Appli
	258	44	1.4	201	10	US-10-995-561-36364 Sequence 36364, A
	259	44	1.4	201	10	US-10-995-561-42346 Sequence 42346, A
	260	44	1.4	201	17	US-11-124-367A-21173 Sequence 21173, A
	261	44	1.4	201	17	US-11-124-367A-32954 Sequence 32954, A
c	262	44	1.4	423	7	US-09-925-065A-566681 Sequence 566681,
c	263	44	1.4	423	7	US-09-925-065A-566682 Sequence 566682,
c	264	44	1.4	448	11	US-10-301-480-15039 Sequence 15039, A
c	265	44	1.4	448	12	US-10-301-480-628448 Sequence 628448,
c	266	44	1.4	464	7	US-09-925-065A-602339 Sequence 602339,
	267	44	1.4	470	7	US-09-925-065A-767971 Sequence 767971,
c	268	44	1.4	496	7	US-09-925-065A-876040 Sequence 876040,
c	269	44	1.4	496	7	US-09-925-065A-876041 Sequence 876041,
c	270	44	1.4	496	7	US-09-925-065A-876042 Sequence 876042,
c	271	44	1.4	496	7	US-09-925-065A-905372 Sequence 905372,
	272	44	1.4	525	12	US-10-301-480-508206 Sequence 508206,
	273	44	1.4	525	12	US-10-301-480-508207 Sequence 508207,
	274	44	1.4	525	12	US-10-301-480-1121615 Sequence 1121615,
	275	44	1.4	525	12	US-10-301-480-1121616 Sequence 1121616,
	276	44	1.4	530	7	US-09-925-065A-297221 Sequence 297221,
	277	44	1.4	530	7	US-09-925-065A-297222 Sequence 297222,
	278	44	1.4	537	7	US-09-925-065A-448591 Sequence 448591,
	279	44	1.4	537	7	US-09-925-065A-448592 Sequence 448592,
	280	44	1.4	539	12	US-10-301-480-611857 Sequence 611857,
	281	44	1.4	539	12	US-10-301-480-1225266 Sequence 1225266,
	282	44	1.4	541	12	US-10-301-480-374012 Sequence 374012,
	283	44	1.4	541	12	US-10-301-480-987421 Sequence 987421,
c	284	44	1.4	556	7	US-09-925-065A-109914 Sequence 109914,
c	285	44	1.4	557	7	US-09-925-065A-870384 Sequence 870384,
c	286	44	1.4	557	7	US-09-925-065A-870385 Sequence 870385,
	287	44	1.4	557	12	US-10-301-480-597680 Sequence 597680,
	288	44	1.4	557	12	US-10-301-480-1211089 Sequence 1211089,
	289	44	1.4	562	7	US-09-925-065A-78244 Sequence 78244, A
	290	44	1.4	562	11	US-10-301-480-179483 Sequence 179483,
	291	44	1.4	562	12	US-10-301-480-792892 Sequence 792892,
c	292	44	1.4	563	11	US-10-301-480-15038 Sequence 15038, A
c	293	44	1.4	563	12	US-10-301-480-628447 Sequence 628447,
c	294	44	1.4	566	11	US-10-301-480-209509 Sequence 209509,
c	295	44	1.4	566	12	US-10-301-480-822918 Sequence 822918,
	296	44	1.4	577	7	US-09-925-065A-707681 Sequence 707681,
	297	44	1.4	577	7	US-09-925-065A-707682 Sequence 707682,
	298	44	1.4	584	12	US-10-301-480-319474 Sequence 319474,
	299	44	1.4	584	12	US-10-301-480-319476 Sequence 319476,
	300	44	1.4	584	12	US-10-301-480-932883 Sequence 932883,
	301	44	1.4	584	12	US-10-301-480-932885 Sequence 932885,
	302	44	1.4	585	12	US-10-301-480-611357 Sequence 611357,
	303	44	1.4	585	12	US-10-301-480-1224766 Sequence 1224766,
	304	44	1.4	588	7	US-09-925-065A-729173 Sequence 729173,
c	305	44	1.4	589	7	US-09-925-065A-914465 Sequence 914465,
c	306	44	1.4	591	7	US-09-925-065A-936386 Sequence 936386,
c	307	44	1.4	593	7	US-09-925-065A-237003 Sequence 237003,
c	308	44	1.4	593	7	US-09-925-065A-237005 Sequence 237005,
c	309	44	1.4	595	7	US-09-925-065A-685243 Sequence 685243,
c	310	44	1.4	599	7	US-09-925-065A-902996 Sequence 902996,
	311	44	1.4	612	7	US-09-925-065A-21972 Sequence 21972, A
	312	44	1.4	612	11	US-10-301-480-123209 Sequence 123209,
	313	44	1.4	612	12	US-10-301-480-736618 Sequence 736618,
c	314	44	1.4	614	11	US-10-301-480-95402 Sequence 95402, A
c	315	44	1.4	614	12	US-10-301-480-708811 Sequence 708811,
	316	44	1.4	616	12	US-10-301-480-534937 Sequence 534937,
	317	44	1.4	616	12	US-10-301-480-1148346 Sequence 1148346,
c	318	44	1.4	618	7	US-09-925-065A-697941 Sequence 697941,
c	319	44	1.4	618	7	US-09-925-065A-697942 Sequence 697942,
c	320	44	1.4	618	7	US-09-925-065A-697943 Sequence 697943,
c	321	44	1.4	619	7	US-09-925-065A-789262 Sequence 789262,
c	322	44	1.4	619	7	US-09-925-065A-789263 Sequence 789263,
c	323	44	1.4	619	7	US-09-925-065A-789264 Sequence 789264,
c	324	44	1.4	619	7	US-09-925-065A-847245 Sequence 847245,
c	325	44	1.4	619	7	US-09-925-065A-951338 Sequence 951338,
	326	44	1.4	623	7	US-09-925-065A-521713 Sequence 521713,
c	327	44	1.4	627	7	US-09-925-065A-545865 Sequence 545865,
c	328	44	1.4	627	7	US-09-925-065A-545866 Sequence 545866,
c	329	44	1.4	627	7	US-09-925-065A-545867 Sequence 545867,
c	330	44	1.4	627	7	US-09-925-065A-923432 Sequence 923432,
	331	44	1.4	629	7	US-09-925-065A-821309 Sequence 821309,
	332	44	1.4	634	11	US-10-301-480-46185 Sequence 46185, A
	333	44	1.4	634	12	US-10-301-480-594891 Sequence 594891,
	334	44	1.4	634	12	US-10-301-480-659594 Sequence 659594,
	335	44	1.4	634	12	US-10-301-480-1208300 Sequence 1208300,
	336	44	1.4	643	7	US-09-925-065A-692017 Sequence 692017,
c	337	44	1.4	659	7	US-09-925-065A-741517 Sequence 741517,
c	338	44	1.4	659	7	US-09-925-065A-741518 Sequence 741518,
	339	44	1.4	668	7	US-09-925-065A-127024 Sequence 127024,
	340	44	1.4	672	12	US-10-301-480-224194 Sequence 224194,
	341	44	1.4	672	12	US-10-301-480-837603 Sequence 837603,
	342	44	1.4	673	7	US-09-925-065A-32538 Sequence 32538, A
	343	44	1.4	673	11	US-10-301-480-133776 Sequence 133776,
	344	44	1.4	673	12	US-10-301-480-747185 Sequence 747185,
c	345	44	1.4	698	11	US-10-301-480-87096 Sequence 87096, A
c	346	44	1.4	698	12	US-10-301-480-700505 Sequence 700505,
c	347	44	1.4	716	12	US-10-301-480-609596 Sequence 609596,
c	348	44	1.4	716	12	US-10-301-480-1223005 Sequence 1223005,
	349	44	1.4	717	12	US-10-301-480-564215 Sequence 564215,
	350	44	1.4	717	12	US-10-301-480-1177624 Sequence 1177624,
c	351	44	1.4	718	12	US-10-301-480-582616 Sequence 582616,
c	352	44	1.4	718	12	US-10-301-480-1196025 Sequence 1196025,
c	353	44	1.4	761	12	US-10-301-480-574634 Sequence 574634,
c	354	44	1.4	761	12	US-10-301-480-1188043 Sequence 1188043,
c	355	44	1.4	783	12	US-10-301-480-582777 Sequence 582777,
c	356	44	1.4	783	12	US-10-301-480-1196186 Sequence 1196186,
c	357	44	1.4	784	7	US-09-925-065A-927249 Sequence 927249,
c	358	44	1.4	822	12	US-10-301-480-574635 Sequence 574635,
c	359	44	1.4	822	12	US-10-301-480-1188044 Sequence 1188044,
c	360	44	1.4	873	11	US-10-301-480-89331 Sequence 89331, A
c	361	44	1.4	873	12	US-10-301-480-702740 Sequence 702740,
c	362	44	1.4	934	12	US-10-301-480-535654 Sequence 535654,
c	363	44	1.4	934	12	US-10-301-480-1149063 Sequence 1149063,
c	364	44	1.4	949	12	US-10-301-480-535653 Sequence 535653,
c	365	44	1.4	949	12	US-10-301-480-1149062 Sequence 1149062,
c	366	44	1.4	955	12	US-10-301-480-571291 Sequence 571291,
c	367	44	1.4	955	12	US-10-301-480-1184700 Sequence 1184700,
c	368	44	1.4	971	12	US-10-301-480-603989 Sequence 603989,
c	369	44	1.4	971	12	US-10-301-480-1217398 Sequence 1217398,
c	370	44	1.4	972	12	US-10-301-480-540105 Sequence 540105,
	371	44	1.4	972	12	US-10-301-480-543642 Sequence 543642,
	372	44	1.4	972	12	US-10-301-480-543643 Sequence 543643,
c	373	44	1.4	972	12	US-10-301-480-1153514 Sequence 1153514,
	374	44	1.4	972	12	US-10-301-480-1157051 Sequence 1157051,
	375	44	1.4	972	12	US-10-301-480-1157052 Sequence 1157052,
	376	44	1.4	979	12	US-10-301-480-587102 Sequence 587102,
	377	44	1.4	979	12	US-10-301-480-1200511 Sequence 1200511,
c	378	44	1.4	980	12	US-10-301-480-606905 Sequence 606905,
c	379	44	1.4	980	12	US-10-301-480-1220314 Sequence 1220314,

c 380	44	1.4	985	12	US-10-301-480-588232	Sequence 588232,
c 381	44	1.4	985	12	US-10-301-480-1201641	Sequence 1201641,
c 382	44	1.4	996	12	US-10-301-480-576981	Sequence 576981,
c 383	44	1.4	996	12	US-10-301-480-1190390	Sequence 1190390,
c 384	44	1.4	998	12	US-10-301-480-543884	Sequence 543884,
c 385	44	1.4	998	12	US-10-301-480-1157293	Sequence 1157293,
c 386	44	1.4	999	12	US-10-301-480-543138	Sequence 543138,
387	44	1.4	999	12	US-10-301-480-563536	Sequence 563536,
c 388	44	1.4	999	12	US-10-301-480-1156547	Sequence 1156547,
389	44	1.4	999	12	US-10-301-480-1176945	Sequence 1176945,
c 390	44	1.4	1000	12	US-10-301-480-585733	Sequence 585733,
c 391	44	1.4	1000	12	US-10-301-480-1199142	Sequence 1199142,
c 392	44	1.4	1005	7	US-09-925-065A-726951	Sequence 726951,
c 393	44	1.4	1005	7	US-09-925-065A-726952	Sequence 726952,
394	44	1.4	1068	7	US-09-925-065A-716304	Sequence 716304,
395	44	1.4	1068	7	US-09-925-065A-716305	Sequence 716305,
396	44	1.4	1068	7	US-09-925-065A-716306	Sequence 716306,
397	44	1.4	1068	7	US-09-925-065A-716307	Sequence 716307,
398	44	1.4	1068	7	US-09-925-065A-716308	Sequence 716308,
c 399	44	1.4	1103	11	US-10-301-480-92758	Sequence 92758, A
c 400	44	1.4	1103	11	US-10-301-480-92759	Sequence 92759, A
c 401	44	1.4	1103	12	US-10-301-480-706167	Sequence 706167,
c 402	44	1.4	1103	12	US-10-301-480-706168	Sequence 706168,
c 403	44	1.4	1122	7	US-09-925-065A-551840	Sequence 551840,
c 404	44	1.4	1122	12	US-10-301-480-529790	Sequence 529790,
c 405	44	1.4	1122	12	US-10-301-480-1143199	Sequence 1143199,
c 406	44	1.4	1141	11	US-10-301-480-40080	Sequence 40080, A
c 407	44	1.4	1141	12	US-10-301-480-653489	Sequence 653489,
408	44	1.4	1228	11	US-10-301-480-34420	Sequence 34420, A
409	44	1.4	1228	12	US-10-301-480-647829	Sequence 647829,
410	44	1.4	1240	11	US-10-301-480-98516	Sequence 98516, A
411	44	1.4	1240	12	US-10-301-480-711925	Sequence 711925,
412	44	1.4	1242	7	US-09-925-065A-712795	Sequence 712795,
413	44	1.4	1330	7	US-09-925-065A-58322	Sequence 58322, A
414	44	1.4	1330	7	US-09-925-065A-58323	Sequence 58323, A
415	44	1.4	1330	7	US-09-925-065A-58324	Sequence 58324, A
416	44	1.4	1330	11	US-10-301-480-159560	Sequence 159560,
417	44	1.4	1330	11	US-10-301-480-159561	Sequence 159561,
418	44	1.4	1330	11	US-10-301-480-159562	Sequence 159562,
419	44	1.4	1330	12	US-10-301-480-772969	Sequence 772969,
420	44	1.4	1330	12	US-10-301-480-772970	Sequence 772970,
421	44	1.4	1330	12	US-10-301-480-772971	Sequence 772971,
c 422	44	1.4	1390	7	US-09-925-065A-47642	Sequence 47642, A
c 423	44	1.4	1390	11	US-10-301-480-148880	Sequence 148880,
c 424	44	1.4	1390	12	US-10-301-480-762289	Sequence 762289,
425	44	1.4	1427	7	US-09-925-065A-77993	Sequence 77993, A
426	44	1.4	1427	11	US-10-301-480-179232	Sequence 179232,
427	44	1.4	1427	12	US-10-301-480-792641	Sequence 792641,
428	44	1.4	1753	7	US-09-925-065A-714611	Sequence 714611,
429	44	1.4	1753	7	US-09-925-065A-714612	Sequence 714612,
430	44	1.4	1753	7	US-09-925-065A-714613	Sequence 714613,
431	44	1.4	1753	7	US-09-925-065A-714614	Sequence 714614,
c 432	44	1.4	1900	11	US-10-301-480-97935	Sequence 97935, A
c 433	44	1.4	1900	12	US-10-301-480-711344	Sequence 711344,
434	44	1.4	1970	7	US-09-925-065A-721388	Sequence 721388,
c 435	44	1.4	2078	7	US-09-925-065A-16796	Sequence 16796, A
c 436	44	1.4	2078	11	US-10-301-480-118033	Sequence 118033,
c 437	44	1.4	2078	12	US-10-301-480-731442	Sequence 731442,
438	44	1.4	2184	7	US-09-925-065A-715885	Sequence 715885,
439	44	1.4	2184	7	US-09-925-065A-715886	Sequence 715886,
440	44	1.4	2184	7	US-09-925-065A-715887	Sequence 715887,
441	44	1.4	2184	7	US-09-925-065A-715888	Sequence 715888,
c 442	44	1.4	2725	7	US-09-925-065A-696477	Sequence 696477,
c 443	44	1.4	3488	7	US-09-925-065A-84868	Sequence 84868, A
c 444	44	1.4	3488	11	US-10-301-480-186108	Sequence 186108,
c 445	44	1.4	3488	12	US-10-301-480-799517	Sequence 799517,
446	44	1.4	23943	11	US-10-330-773-919	Sequence 919, App
447	44	1.4	67467	17	US-11-124-368A-2889	Sequence 2889, Ap
448	44	1.4	86361	10	US-10-995-561-13364	Sequence 13364, A
c 449	44	1.4	88116	10	US-10-995-561-13351	Sequence 13351, A
c 450	44	1.4	110847	17	US-11-121-086-11	Sequence 11, Appl
c 451	44	1.4	138941	9	US-10-489-730-10	GENERAL INFORMATI
c 452	44	1.4	139573	12	US-10-506-513-3	Sequence 3, Appli
453	44	1.4	150468	17	US-11-112-908-56	Sequence 56, Appl
c 454	44	1.4	159440	13	US-10-960-414-12	Sequence 12, Appl
455	44	1.4	168656	17	US-11-112-908-59	Sequence 59, Appl
c 456	44	1.4	168753	17	US-11-181-234-1	Sequence 1, Appli
457	44	1.4	170285	17	US-11-112-908-58	Sequence 58, Appl
c 458	44	1.4	171162	17	US-11-112-908-38	Sequence 38, Appl
459	44	1.4	173602	17	US-11-121-086-25	Sequence 25, Appl
460	44	1.4	188056	17	US-11-120-925-1	Sequence 1, Appli
461	44	1.4	193789	17	US-11-112-908-55	Sequence 55, Appl
c 462	44	1.4	199130	10	US-10-995-561-13233	Sequence 13233, A
463	44	1.4	220895	10	US-10-775-169-88	Sequence 88, Appl
464	44	1.4	222094	10	US-10-995-561-13244	Sequence 13244, A
465	44	1.4	1125000	10	US-10-995-561-13286	Sequence 13286, A
466	43	1.4	201	10	US-10-995-561-28594	Sequence 28594, A
467	43	1.4	201	10	US-10-995-561-42347	Sequence 42347, A
c 468	43	1.4	201	10	US-10-995-561-49969	Sequence 49969, A
c 469	43	1.4	201	10	US-10-995-561-50808	Sequence 50808, A
c 470	43	1.4	201	10	US-10-995-561-61087	Sequence 61087, A
471	43	1.4	201	10	US-10-995-561-68769	Sequence 68769, A
c 472	43	1.4	201	10	US-10-995-561-80611	Sequence 80611, A
c 473	43	1.4	201	10	US-10-995-561-80638	Sequence 80638, A
474	43	1.4	201	17	US-11-124-367A-32955	Sequence 32955, A
c 475	43	1.4	444	7	US-09-925-065A-932412	Sequence 932412,
c 476	43	1.4	444	7	US-09-925-065A-951567	Sequence 951567,
c 477	43	1.4	444	7	US-09-925-065A-951568	Sequence 951568,
c 478	43	1.4	462	11	US-10-301-480-5149	Sequence 5149, Ap
c 479	43	1.4	462	12	US-10-301-480-618558	Sequence 618558,
c 480	43	1.4	463	7	US-09-925-065A-501502	Sequence 501502,
481	43	1.4	472	7	US-09-925-065A-815432	Sequence 815432,
c 482	43	1.4	486	7	US-09-925-065A-168216	Sequence 168216,
483	43	1.4	528	7	US-09-925-065A-491876	Sequence 491876,
484	43	1.4	528	7	US-09-925-065A-491877	Sequence 491877,
485	43	1.4	539	7	US-09-925-065A-485618	Sequence 485618,
c 486	43	1.4	542	7	US-09-925-065A-411190	Sequence 411190,
c 487	43	1.4	553	7	US-09-925-065A-908993	Sequence 908993,
488	43	1.4	559	7	US-09-925-065A-651202	Sequence 651202,
489	43	1.4	567	7	US-09-925-065A-581250	Sequence 581250,
490	43	1.4	570	7	US-09-925-065A-586025	Sequence 586025,
491	43	1.4	572	7	US-09-925-065A-11001	Sequence 11001, A
c 492	43	1.4	572	7	US-09-925-065A-516598	Sequence 516598,
493	43	1.4	572	11	US-10-301-480-112238	Sequence 112238,
494	43	1.4	572	12	US-10-301-480-725647	Sequence 725647,
c 495	43	1.4	577	11	US-10-301-480-45127	Sequence 45127, A
c 496	43	1.4	577	12	US-10-301-480-658536	Sequence 658536,
c 497	43	1.4	581	7	US-09-925-065A-926250	Sequence 926250,
c 498	43	1.4	590	7	US-09-925-065A-762418	Sequence 762418,
c 499	43	1.4	590	7	US-09-925-065A-762419	Sequence 762419,
c 500	43	1.4	593	7	US-09-925-065A-411191	Sequence 411191,

ALIGNMENTS

RESULT 1
US-09-925-065A-740956/c
; Sequence 740956, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 740956
;.  LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-740956

  Query Match             16.3%;  Score 509;  DB 7;  Length 560;
  Best Local Similarity   99.8%;  Pred. No. 9.1e-70;
  Matches  559;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        199 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 CATCACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCC 501

Qy        259 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CATCCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGG 441

Qy        319 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCGAAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTA 381

Qy        379 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAA 438
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        380 CCACTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAA 321

Qy        439 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 ACCCTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTT 261

Qy        499 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GCCCCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCAC 201

Qy        559 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CGTGGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGAT 141

Qy        619 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GGACGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCG 81

Qy        679 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CACGTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGG 21

Qy        739 ACCATGAAGAGCGTTCGTGC 758
              ||||||||||||||||||||
Db         20 ACCATGAAGAGCGTTCGTGC 1


RESULT 2
US-09-925-065A-769508/c
; Sequence 769508, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 769508
;   LENGTH: 557
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-769508

  Query Match             16.2%;  Score 506;  DB 7;  Length 557;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-69;
  Matches  556;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        202 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 CACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCAT 498

Qy        262 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CCGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCG 438

Qy        322 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 AAGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCA 378

Qy        382 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACC 441
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        377 CTGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACC 318

Qy        442 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CTACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCC 258

Qy        502 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGT 198

Qy        562 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GGAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGA 138

Qy        622 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 CGAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCAC 78

Qy        682 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GTCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACC 18

Qy        742 ATGAAGAGCGTTCGTGC 758
              |||||||||||||||||
Db         17 ATGAAGAGCGTTCGTGC 1


RESULT 3
US-09-925-065A-736351/c
; Sequence 736351, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

```
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 736351
;   LENGTH: 556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-736351

  Query Match             16.2%;  Score 505;  DB 7;  Length 556;
  Best Local Similarity   99.8%;  Pred. No. 3.7e-69;
  Matches  555;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        203 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 ACAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATC 497

Qy        263 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CGTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGA 437

Qy        323 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCAC 377

Qy        383 TGATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCC 442
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        376 TGATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCC 317

Qy        443 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCC 257

Qy        503 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CGGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTG 197

Qy        563 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GAGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGAC 137

Qy        623 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GAGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACG 77

Qy        683 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCA 17

Qy        743 TGAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db         16 TGAAGAGCGTTCGTGC 1


RESULT 4
US-09-925-065A-737120
; Sequence 737120, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;, TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737120
;   LENGTH: 555
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-737120

  Query Match             16.1%;  Score 504;  DB 7;  Length 555;
  Best Local Similarity   99.8%;  Pred. No. 5.3e-69;
  Matches  554;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        204 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGCAAAAACTCCGCGGGAGCCTGGCCCGCTTTTTACCTGGGCCTCAGTTTCCCCATCC 60

Qy        264 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTAAAATAGAACGGGTTGGATCTCCCGAGCGCTAACATTCCAGAACTCGGATGGGGCGAA 120

Qy        324 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGGGAGGGAGGGATGGGCCACCCACACGTGACCTCCCCGCGTGGAGCCCCGCCTACCACT 180

Qy        384 GATCCAGGGGGTGGCAGCTCCGGCCGGGACGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 443
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        181 GATCCAGGGGGTGGCAGCTCCGGCCGGGAYGAGCGGGGTGGGCGGGTCCTAGGAAACCCT 240

Qy        444 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCGGCCGCCCTTGGCAGCGCCTAAGGCGGAGCGCGCGGCTCTGCAGCCTGCTTGCCCC 300

Qy        504 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGTTGGCACCCACGGAGGATGGGGACCGCACCCTCAGCTTCGCAGGGAGCCACCGTGG 360

Qy        564 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCAGGGCGGTGCAGAGACACGACGTGTGACTCGGAGTGCGCCTGGGGAGGATGGACG 420

Qy        624 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGGAGCGGGGGACCGCTAACGGGGCTCCCTCTGCGCGCCCCGTCCGCAGAGGCGCACGT 480

Qy        684 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAGGGTCCCGGGCGGGCTCCGTGGACGTTGGCGGTAGCGCCGAGCGAGTCACGGACCAT 540

Qy        744 GAAGAGCGTTCGTGC 758
              ||||||||||||||||
Db        541 GAAGAGCGTTCGTGC 555


RESULT 5
US-10-301-480-20199/c
; Sequence 20199, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20199
;   LENGTH: 477
;   TYPE: DNA
;.  ORGANISM: Homo sapien
US-10-301-480-20199

  Query Match             1.7%;  Score 53;  DB 11;  Length 477;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 50


RESULT 6
US-10-301-480-633608/c
; Sequence 633608, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633608
;   LENGTH: 477
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-633608

  Query Match             1.7%;  Score 53;  DB 12;  Length 477;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 50


RESULT 7
US-10-301-480-20198/c
; Sequence 20198, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20198
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-20198

  Query Match             1.7%;  Score 53;  DB 11;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 109


RESULT 8
US-10-301-480-633607/c
; Sequence 633607, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633607
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-633607

  Query Match             1.7%;  Score 53;  DB 12;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 109


RESULT 9
US-09-925-065A-591947/c
; Sequence 591947, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591947
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591947

  Query Match             1.7%;  Score 52;  DB 7;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 113


RESULT 10
US-09-925-065A-768118
; Sequence 768118, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 768118
;   LENGTH: 574
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-768118

  Query Match             1.7%;  Score 52;  DB 7;  Length 574;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 286


RESULT 11
US-09-925-065A-59682/c
; Sequence 59682, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59682
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-59682

  Query Match             1.7%;  Score 52;  DB 7;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 124


RESULT 12
US-10-301-480-160920/c
; Sequence 160920, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 160920
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-160920

  Query Match             1.7%;  Score 52;  DB 11;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 124


RESULT 13
US-10-301-480-774329/c
; Sequence 774329, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 774329
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-774329

  Query Match             1.7%;  Score 52;  DB 12;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 124


RESULT 14
US-10-301-480-454942/c
```

```
; Sequence 454942, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 454942
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-454942

  Query Match             1.7%;  Score 52;  DB 12;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255


RESULT 15
US-10-301-480-1068351/c
; Sequence 1068351, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1068351
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1068351

  Query Match             1.7%;  Score 52;  DB 12;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255


RESULT 16
US-09-925-065A-386361/c
; Sequence 386361, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 386361
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-386361

  Query Match             1.7%;  Score 52;  DB 7;  Length 617;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255


RESULT 17
US-10-301-480-611724/c
; Sequence 611724, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 611724
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-611724

  Query Match             1.7%;  Score 52;  DB 12;  Length 618;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 548


RESULT 18
US-10-301-480-1225133/c
; Sequence 1225133, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
```

```
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1225133
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1225133

  Query Match             1.7%;  Score 52;  DB 12;  Length 618;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 548


RESULT 19
US-09-925-065A-708701/c
; Sequence 708701, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 708701
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-708701

  Query Match             1.7%;  Score 52;  DB 7;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 439


RESULT 20
US-10-301-480-598972
; Sequence 598972, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 598972
;   LENGTH: 974
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-598972

  Query Match             1.7%;  Score 52;  DB 12;  Length 974;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 967


RESULT 21
US-10-301-480-1212381
; Sequence 1212381, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1212381
;   LENGTH: 974
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1212381

  Query Match             1.7%;  Score 52;  DB 12;  Length 974;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2889 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 967


RESULT 22
US-11-121-086-84
; Sequence 84, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 84
;   LENGTH: 118996
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-84

  Query Match             1.7%;  Score 52;  DB 17;  Length 118996;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73872 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 73923


RESULT 23
US-11-121-086-47
; Sequence 47, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 47
;   LENGTH: 137671
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-47

  Query Match             1.7%;  Score 52;  DB 17;  Length 137671;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75583 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 75634


RESULT 24
US-10-995-561-13396/c
; Sequence 13396, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13396
;   LENGTH: 398287
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(398287)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13396

  Query Match             1.7%;  Score 52;  DB 10;  Length 398287;
  Best Local Similarity   100.0%;  Pred. No. 0.09;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281811 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 281760


RESULT 25
US-09-925-065A-211249/c
; Sequence 211249, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 211249
;   LENGTH: 385
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-211249

  Query Match             1.6%;  Score 51;  DB 7;  Length 385;
  Best Local Similarity   100.0%;  Pred. No. 4.5;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 117


RESULT 26
US-09-925-065A-141833
; Sequence 141833, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 141833
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-141833

  Query Match             1.6%;  Score 51;  DB 7;  Length 561;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 452


RESULT 27
US-10-301-480-236672
; Sequence 236672, Application US/10301480
```

```
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 236672
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-236672

  Query Match             1.6%;  Score 51;  DB 12;  Length 561;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 452


RESULT 28
US-10-301-480-850081
; Sequence 850081, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 850081
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-850081

  Query Match             1.6%;  Score 51;  DB 12;  Length 561;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 452


RESULT 29
US-10-301-480-1011
; Sequence 1011, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1011
;   LENGTH: 587
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1011

  Query Match             1.6%;  Score 51;  DB 11;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 417


RESULT 30
US-10-301-480-42180
; Sequence 42180, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42180
;   LENGTH: 587
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-42180

  Query Match             1.6%;  Score 51;  DB 11;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 417


RESULT 31
US-10-301-480-614420
; Sequence 614420, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 614420
;   LENGTH: 587
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

US-10-301-480-614420

Query Match 1.6%; Score 51; DB 12; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 367 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 417

RESULT 32
US-10-301-480-655589
; Sequence 655589, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 655589
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-655589

Query Match 1.6%; Score 51; DB 12; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 367 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 417

RESULT 33
US-10-301-480-589593/c
; Sequence 589593, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 589593
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-589593

Query Match 1.6%; Score 51; DB 12; Length 852;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 494 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 444

RESULT 34
US-10-301-480-1203002/c
; Sequence 1203002, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1203002
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1203002

Query Match 1.6%; Score 51; DB 12; Length 852;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 494 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 444

RESULT 35
US-11-121-086-107/c
; Sequence 107, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 107
; LENGTH: 197096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-107

Query Match 1.6%; Score 51; DB 17; Length 197096;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 68932 AGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 68882

RESULT 36
US-09-925-065A-484883/c
; Sequence 484883, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135

```
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 484883
;   LENGTH: 571
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-484883

  Query Match             1.6%;  Score 50;  DB 7;  Length 571;
  Best Local Similarity   100.0%;  Pred. No. 5.2;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 222


RESULT 37
US-09-925-065A-785967/c
; Sequence 785967, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785967
;   LENGTH: 572
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-785967

  Query Match             1.6%;  Score 50;  DB 7;  Length 572;
  Best Local Similarity   100.0%;  Pred. No. 5.2;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 298


RESULT 38
US-09-925-065A-894777/c
; Sequence 894777, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 894777
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-894777

  Query Match             1.6%;  Score 50;  DB 7;  Length 625;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 180


RESULT 39
US-09-925-065A-892664/c
; Sequence 892664, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892664
;   LENGTH: 630
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892664

  Query Match             1.6%;  Score 50;  DB 7;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 185


RESULT 40
US-09-925-065A-892405/c
; Sequence 892405, Application US/09925065A
```

```
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892405
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892405

  Query Match             1.6%;  Score 50;  DB 7;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 188


RESULT 41
US-09-925-065A-892514/c
; Sequence 892514, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 892514
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-892514

  Query Match             1.6%;  Score 50;  DB 7;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 188


RESULT 42
US-09-925-065A-949927
; Sequence 949927, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 949927
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-949927

  Query Match             1.6%;  Score 50;  DB 7;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 4.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 78


RESULT 43
US-09-925-065A-917225
; Sequence 917225, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 917225
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-917225

  Query Match             1.6%;  Score 50;  DB 7;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497

RESULT 44
US-09-925-065A-917226
; Sequence 917226, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 917226
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-917226

  Query Match             1.6%;  Score 50;  DB 7;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497

RESULT 45
US-09-925-065A-944569
; Sequence 944569, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944569
;   LENGTH: 676
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944569

  Query Match             1.6%;  Score 50;  DB 7;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 497


RESULT 46
US-09-925-065A-934608
; Sequence 934608, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 934608
;   LENGTH: 739
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-934608

  Query Match             1.6%;  Score 50;  DB 7;  Length 739;
  Best Local Similarity   100.0%;  Pred. No. 4.5;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 652


RESULT 47
US-09-925-065A-941367
; Sequence 941367, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 941367
;   LENGTH: 742
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-941367
```

```
  Query Match             1.6%;  Score 50;  DB 7;  Length 742;
  Best Local Similarity   100.0%;  Pred. No. 4.5;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 652


RESULT 48
US-10-301-480-567391
; Sequence 567391, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 567391
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-567391

  Query Match             1.6%;  Score 50;  DB 12;  Length 999;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 156


RESULT 49
US-10-301-480-1180800
; Sequence 1180800, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1180800
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1180800

  Query Match             1.6%;  Score 50;  DB 12;  Length 999;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 156


RESULT 50
US-10-301-480-100096
; Sequence 100096, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100096
;   LENGTH: 1986
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-100096

  Query Match             1.6%;  Score 50;  DB 11;  Length 1986;
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 246


RESULT 51
US-10-301-480-713505
; Sequence 713505, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 713505
;   LENGTH: 1986
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-713505

  Query Match             1.6%;  Score 50;  DB 12;  Length 1986;
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 246


RESULT 52
US-11-121-086-83/c
; Sequence 83, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
```

; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 83
; LENGTH: 187745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-83

Query Match 1.6%; Score 50; DB 17; Length 187745;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159390 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 159341

RESULT 53
US-11-112-908-21
; Sequence 21, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 207908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-21

Query Match 1.6%; Score 50; DB 17; Length 207908;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 2937
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168125 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 168174

RESULT 54
US-11-112-908-19
; Sequence 19, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19
; LENGTH: 212805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-19

Query Match 1.6%; Score 50; DB 17; Length 212805;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 2937
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24046 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAA 24095

RESULT 55
US-10-301-480-235645/c
; Sequence 235645, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 235645
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-235645

Query Match 1.6%; Score 49; DB 12; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 273 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225

RESULT 56
US-10-301-480-849054/c
; Sequence 849054, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 849054
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien

```
US-10-301-480-849054

  Query Match             1.6%;  Score 49;  DB 12;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 8.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225


RESULT 57
US-09-925-065A-140670/c
; Sequence 140670, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 140670
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-140670

  Query Match             1.6%;  Score 49;  DB 7;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 8.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225


RESULT 58
US-10-301-480-15264/c
; Sequence 15264, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15264
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-15264

  Query Match             1.6%;  Score 49;  DB 11;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182


RESULT 59
US-10-301-480-235646/c
; Sequence 235646, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 235646
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-235646

  Query Match             1.6%;  Score 49;  DB 12;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182


RESULT 60
US-10-301-480-628673/c
; Sequence 628673, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 628673
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-628673

  Query Match             1.6%;  Score 49;  DB 12;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182


RESULT 61
```

```
US-10-301-480-849055/c
; Sequence 849055, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 849055
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-849055

  Query Match             1.6%;  Score 49;  DB 12;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182


RESULT 62
US-09-925-065A-140671/c
; Sequence 140671, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 140671
;   LENGTH: 532
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-140671

  Query Match             1.6%;  Score 49;  DB 7;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 184


RESULT 63
US-09-925-065A-18609/c
; Sequence 18609, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18609
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18609

  Query Match             1.6%;  Score 49;  DB 7;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 64
US-09-925-065A-18610/c
; Sequence 18610, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18610
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18610

  Query Match             1.6%;  Score 49;  DB 7;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 65
```

```
US-09-925-065A-18611/c
; Sequence 18611, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18611
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-18611

  Query Match             1.6%;  Score 49;  DB 7;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 66
US-10-301-480-119846/c
; Sequence 119846, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119846
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-119846

  Query Match             1.6%;  Score 49;  DB 11;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 67
US-10-301-480-119847/c
; Sequence 119847, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119847
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-119847

  Query Match             1.6%;  Score 49;  DB 11;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 68
US-10-301-480-119848/c
; Sequence 119848, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119848
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-119848

  Query Match             1.6%;  Score 49;  DB 11;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 69
US-10-301-480-733255/c
; Sequence 733255, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
```

```
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 733255
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-733255

  Query Match             1.6%;  Score 49;  DB 12;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 70
US-10-301-480-733256/c
; Sequence 733256, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 733256
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-733256

  Query Match             1.6%;  Score 49;  DB 12;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 71
US-10-301-480-733257/c
; Sequence 733257, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 733257
;   LENGTH: 552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-733257

  Query Match             1.6%;  Score 49;  DB 12;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 315


RESULT 72
US-09-925-065A-708700/c
; Sequence 708700, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 708700
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-708700

  Query Match             1.6%;  Score 49;  DB 7;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 6.6;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2892 GCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 439


RESULT 73
US-09-925-065A-932035/c
; Sequence 932035, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932035
;   LENGTH: 750
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932035

  Query Match             1.6%;  Score 49;  DB 7;  Length 750;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 564


RESULT 74
US-09-925-065A-932036/c
; Sequence 932036, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 932036
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-932036

  Query Match             1.6%;  Score 49;  DB 7;  Length 750;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 564


RESULT 75
US-10-508-307-14/c
; Sequence 14, Application US/10508307
; Publication No. US20060084790A1
; GENERAL INFORMATION:
;  APPLICANT: Stephens, Len
;  APPLICANT:  Hawkins, Phillip Thomas
;  TITLE OF INVENTION: P-REX1, A PTDINS (3,4,5) P3-G-BETA-GAMMA-
;  TITLE OF INVENTION:  REGULATED GUANINE-NUCLEOTIDE EXCHANGE FACTOR
;  TITLE OF INVENTION:  FOR RAC
;  FILE REFERENCE: 130217.401USPC
;  CURRENT APPLICATION NUMBER: US/10/508,307
;  CURRENT FILING DATE:  2004-09-21
;  PRIOR APPLICATION NUMBER: PCT/GB03/01238
;  PRIOR FILING DATE: 2003-03-21
;  PRIOR APPLICATION NUMBER: GB 0206684.3
;  PRIOR FILING DATE: 2002-03-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 243428
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-508-307-14

  Query Match             1.6%;  Score 49;  DB 13;  Length 243428;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3074 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db     235339 GATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 235291


RESULT 76
US-10-301-480-24305
; Sequence 24305, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24305
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-24305

  Query Match             1.5%;  Score 48;  DB 11;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 208


RESULT 77
US-10-301-480-637714
; Sequence 637714, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 637714
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-637714

  Query Match             1.5%;  Score 48;  DB 12;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 208


RESULT 78
US-10-301-480-43148
; Sequence 43148, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43148
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-43148

  Query Match             1.5%;  Score 48;  DB 11;  Length 438;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 208


RESULT 79
US-10-301-480-656557
; Sequence 656557, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 656557
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-656557

  Query Match             1.5%;  Score 48;  DB 12;  Length 438;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 208


RESULT 80
US-10-301-480-597620/c
; Sequence 597620, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 597620
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-597620

  Query Match             1.5%;  Score 48;  DB 12;  Length 617;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 332


RESULT 81
US-10-301-480-1211029/c
; Sequence 1211029, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1211029
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1211029

  Query Match             1.5%;  Score 48;  DB 12;  Length 617;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 332


RESULT 82
US-10-301-480-597621/c
; Sequence 597621, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
```

```
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 597621
;    LENGTH: 802
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-597621

  Query Match             1.5%;  Score 48;  DB 12;  Length 802;
  Best Local Similarity   100.0%;  Pred. No. 8.8;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 332


RESULT 83
US-10-301-480-1211030/c
; Sequence 1211030, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1211030
;    LENGTH: 802
;    TYPE: DNA
;    ORGANISM: Homo sapien
US-10-301-480-1211030

  Query Match             1.5%;  Score 48;  DB 12;  Length 802;
  Best Local Similarity   100.0%;  Pred. No. 8.8;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 332


RESULT 84
US-11-124-368A-2878
; Sequence 2878, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2878
;    LENGTH: 19675
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-11-124-368A-2878

  Query Match             1.5%;  Score 48;  DB 17;  Length 19675;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      14998 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTC 15045


RESULT 85
US-09-925-065A-440136/c
; Sequence 440136, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 440136
;    LENGTH: 534
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-440136

  Query Match             1.5%;  Score 47;  DB 7;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        161 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 115


RESULT 86
US-09-925-065A-440137/c
; Sequence 440137, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 440137
;    LENGTH: 534
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-440137

  Query Match             1.5%;  Score 47;  DB 7;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 3117
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCT 115


RESULT 87
US-09-925-065A-623360/c
; Sequence 623360, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 623360
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-623360

  Query Match             1.5%;  Score 47;  DB 7;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3076 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 TTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 290


RESULT 88
US-10-995-561-64399
; Sequence 64399, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64399
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-64399

  Query Match             1.5%;  Score 46;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db         44 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 89


RESULT 89
US-10-301-480-20200/c
; Sequence 20200, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20200
;   LENGTH: 496
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-20200

  Query Match             1.5%;  Score 46;  DB 11;  Length 496;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 109


RESULT 90
US-10-301-480-633609/c
; Sequence 633609, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 633609
;   LENGTH: 496
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-633609

  Query Match             1.5%;  Score 46;  DB 12;  Length 496;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 109


RESULT 91
US-10-301-480-570118/c
```

```
; Sequence 570118, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 570118
;   LENGTH: 514
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-570118

  Query Match             1.5%;  Score 46;  DB 12;  Length 514;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 298


RESULT 92
US-10-301-480-1183527/c
; Sequence 1183527, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1183527
;   LENGTH: 514
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1183527

  Query Match             1.5%;  Score 46;  DB 12;  Length 514;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 298


RESULT 93
US-10-301-480-577517/c
; Sequence 577517, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577517
;   LENGTH: 540
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-577517

  Query Match             1.5%;  Score 46;  DB 12;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        356 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 311


RESULT 94
US-10-301-480-1190926/c
; Sequence 1190926, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1190926
;   LENGTH: 540
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1190926

  Query Match             1.5%;  Score 46;  DB 12;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        356 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 311


RESULT 95
US-09-925-065A-124249
; Sequence 124249, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 124249
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-124249

  Query Match             1.5%;  Score 46;  DB 7;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 403


RESULT 96
US-10-301-480-221987/c
; Sequence 221987, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 221987
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-221987

  Query Match             1.5%;  Score 46;  DB 12;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 167


RESULT 97
US-10-301-480-835396/c
; Sequence 835396, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 835396
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-835396

  Query Match             1.5%;  Score 46;  DB 12;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 167


RESULT 98
US-09-925-065A-843811/c
; Sequence 843811, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 843811
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-843811

  Query Match             1.5%;  Score 46;  DB 7;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 253


RESULT 99
US-09-925-065A-801712/c
; Sequence 801712, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 801712
;   LENGTH: 609
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-925-065A-801712

  Query Match             1.5%;  Score 46;  DB 7;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 254


RESULT 100
US-09-925-065A-48663/c
; Sequence 48663, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48663
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-48663

  Query Match             1.5%;  Score 46;  DB 7;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 495


RESULT 101
US-09-925-065A-65044
; Sequence 65044, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65044
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-65044

  Query Match             1.5%;  Score 46;  DB 7;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 285


RESULT 102
US-10-301-480-149901/c
; Sequence 149901, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 149901
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-149901

  Query Match             1.5%;  Score 46;  DB 11;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 495


RESULT 103
US-10-301-480-166282
; Sequence 166282, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 166282
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-166282

  Query Match             1.5%;  Score 46;  DB 11;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 285


RESULT 104
US-10-301-480-763310/c
; Sequence 763310, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 763310
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-763310

  Query Match             1.5%;  Score 46;  DB 12;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 495


RESULT 105
US-10-301-480-779691
; Sequence 779691, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 779691
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-779691

  Query Match             1.5%;  Score 46;  DB 12;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 285


RESULT 106
US-09-925-065A-47410/c
; Sequence 47410, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47410
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-47410

  Query Match             1.5%;  Score 46;  DB 7;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 107
US-09-925-065A-47411/c
; Sequence 47411, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47411
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-47411

  Query Match             1.5%;  Score 46;  DB 7;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219
```

```
RESULT 108
US-10-301-480-148648/c
; Sequence 148648, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148648
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-148648

  Query Match             1.5%;  Score 46;  DB 11;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 109
US-10-301-480-148649/c
; Sequence 148649, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148649
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-148649

  Query Match             1.5%;  Score 46;  DB 11;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 110
US-10-301-480-762057/c
; Sequence 762057, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 762057
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-762057

  Query Match             1.5%;  Score 46;  DB 12;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 111
US-10-301-480-762058/c
; Sequence 762058, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 762058
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-762058

  Query Match             1.5%;  Score 46;  DB 12;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 219


RESULT 112
US-09-925-065A-930543/c
; Sequence 930543, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
```

; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 930543
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930543

Query Match 1.5%; Score 46; DB 7; Length 615;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||||
Db 290 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 245

RESULT 113
US-09-925-065A-930544/c
; Sequence 930544, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 930544
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930544

Query Match 1.5%; Score 46; DB 7; Length 615;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
||||||||||||||||||||||||||||||||||||||||||||||
Db 290 TGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 245

RESULT 114
US-09-925-065A-870312/c
; Sequence 870312, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 870312
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-870312

Query Match 1.5%; Score 46; DB 7; Length 616;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
||||||||||||||||||||||||||||||||||||||||||||||
Db 246 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 201

RESULT 115
US-09-925-065A-554324
; Sequence 554324, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 554324
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-554324

Query Match 1.5%; Score 46; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
||||||||||||||||||||||||||||||||||||||||||||
Db 154 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 199

RESULT 116
US-09-925-065A-759345/c
; Sequence 759345, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 759345
;    LENGTH: 654
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-759345

  Query Match             1.5%;  Score 46;  DB 7;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        274 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 229


RESULT 117
US-09-925-065A-790726
; Sequence 790726, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790726
;    LENGTH: 659
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-790726

  Query Match             1.5%;  Score 46;  DB 7;  Length 659;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 225


RESULT 118
US-09-925-065A-790727
; Sequence 790727, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790727
;    LENGTH: 659
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-790727

  Query Match             1.5%;  Score 46;  DB 7;  Length 659;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 225


RESULT 119
US-09-925-065A-771454
; Sequence 771454, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 771454
;    LENGTH: 661
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-771454

  Query Match             1.5%;  Score 46;  DB 7;  Length 661;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||
Db        388 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 433


RESULT 120
US-09-925-065A-737520
; Sequence 737520, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737520
;   LENGTH: 672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-737520

  Query Match             1.5%;  Score 46;  DB 7;  Length 672;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        385 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 430


RESULT 121
US-09-925-065A-737227/c
; Sequence 737227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 737227
;   LENGTH: 680
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-737227

  Query Match             1.5%;  Score 46;  DB 7;  Length 680;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249


RESULT 122
US-09-925-065A-852555
; Sequence 852555, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 852555
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-852555

  Query Match             1.5%;  Score 46;  DB 7;  Length 689;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||
Db        293 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 338


RESULT 123
US-09-925-065A-852556
; Sequence 852556, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 852556
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-852556

  Query Match             1.5%;  Score 46;  DB 7;  Length 689;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||
Db        293 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 338
```

```
RESULT 124
US-09-925-065A-938958/c
; Sequence 938958, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 938958
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938958

  Query Match             1.5%;  Score 46;  DB 7;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 125
US-09-925-065A-938959/c
; Sequence 938959, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 938959
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938959

  Query Match             1.5%;  Score 46;  DB 7;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 126
US-09-925-065A-938960/c
; Sequence 938960, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 938960
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-938960

  Query Match             1.5%;  Score 46;  DB 7;  Length 840;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 101


RESULT 127
US-10-301-480-601877/c
; Sequence 601877, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 601877
;   LENGTH: 989
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-601877

  Query Match             1.5%;  Score 46;  DB 12;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 807
```

```
RESULT 128
US-10-301-480-1215286/c
; Sequence 1215286, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1215286
;   LENGTH: 989
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1215286

  Query Match             1.5%;  Score 46;  DB 12;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        852 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 807


RESULT 129
US-10-301-480-612422/c
; Sequence 612422, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 612422
;   LENGTH: 998
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-612422

  Query Match             1.5%;  Score 46;  DB 12;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        848 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 803


RESULT 130
US-10-301-480-1225831/c
; Sequence 1225831, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1225831
;   LENGTH: 998
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1225831

  Query Match             1.5%;  Score 46;  DB 12;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        848 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 803


RESULT 131
US-09-925-065A-711495
; Sequence 711495, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 711495
;   LENGTH: 1105
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-711495

  Query Match             1.5%;  Score 46;  DB 7;  Length 1105;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 919


RESULT 132
US-09-925-065A-678054
; Sequence 678054, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 678054
;   LENGTH: 1364
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-678054

  Query Match             1.5%;  Score 46;  DB 7;  Length 1364;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3073 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 3118
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTG 396


RESULT 133
US-09-925-065A-689300
; Sequence 689300, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 689300
;   LENGTH: 1635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-689300

  Query Match             1.5%;  Score 46;  DB 7;  Length 1635;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        831 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 876


RESULT 134
US-09-925-065A-677066
; Sequence 677066, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 677066
;   LENGTH: 2227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-677066

  Query Match             1.5%;  Score 46;  DB 7;  Length 2227;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2080


RESULT 135
US-10-301-480-94903
; Sequence 94903, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94903
;   LENGTH: 2798
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-94903

  Query Match             1.5%;  Score 46;  DB 11;  Length 2798;
  Best Local Similarity   100.0%;  Pred. No. 9.3;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 1479


RESULT 136
US-10-301-480-708312
; Sequence 708312, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
```

```
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 708312
;   LENGTH: 2798
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-708312

  Query Match             1.5%;  Score 46;  DB 12;  Length 2798;
  Best Local Similarity   100.0%;  Pred. No. 9.3;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db       1434 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 1479


RESULT 137
US-10-995-561-13369
; Sequence 13369, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13369
;   LENGTH: 56448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(56448)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13369

  Query Match             1.5%;  Score 46;  DB 10;  Length 56448;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db      10994 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 11039


RESULT 138
US-11-112-908-36/c
; Sequence 36, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
;   LENGTH: 98345
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (49350)..(49350)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-112-908-36

  Query Match             1.5%;  Score 46;  DB 17;  Length 98345;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db      18440 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 18395


RESULT 139
US-10-330-773-795
; Sequence 795, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 795
;   LENGTH: 109661
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(109661)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-795

  Query Match             1.5%;  Score 46;  DB 11;  Length 109661;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db      40304 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 40349


RESULT 140
US-11-112-908-54/c
; Sequence 54, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 54
;   LENGTH: 150450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-112-908-54

  Query Match             1.5%;  Score 46;  DB 17;  Length 150450;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db      99596 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 99551


RESULT 141
US-11-112-908-53/c
; Sequence 53, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
;   LENGTH: 191343
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-112-908-53

  Query Match             1.5%;  Score 46;  DB 17;  Length 191343;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
             |||||||||||||||||||||||||||||||||||||||||||||
Db      18838 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 18793


RESULT 142
US-10-995-561-58923/c
; Sequence 58923, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58923
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-58923

  Query Match             1.4%;  Score 45;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        149 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105


RESULT 143
US-10-995-561-58959/c
; Sequence 58959, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58959
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-58959

  Query Match             1.4%;  Score 45;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        148 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 104


RESULT 144
US-10-995-561-68537
; Sequence 68537, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68537
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-68537

  Query Match             1.4%;  Score 45;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2898 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 2942
             |||||||||||||||||||||||||||||||||||||||||||||
Db        128 GGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAG 172


RESULT 145
US-10-301-480-49738/c
; Sequence 49738, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
```

```
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49738
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-49738

  Query Match             1.4%;  Score 45;  DB 11;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
pb        148 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 104


RESULT 146
US-10-301-480-49739/c
; Sequence 49739, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49739
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-49739

  Query Match             1.4%;  Score 45;  DB 11;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 104


RESULT 147
US-10-301-480-663147/c
; Sequence 663147, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;. FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 663147
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-663147

  Query Match             1.4%;  Score 45;  DB 12;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 104


RESULT 148
US-10-301-480-663148/c
; Sequence 663148, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 663148
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-663148

  Query Match             1.4%;  Score 45;  DB 12;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2888 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 2932
             |||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGAGGCAGGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTG 104


RESULT 149
US-10-301-480-9096
; Sequence 9096, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9096
;   LENGTH: 527
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-9096

  Query Match             1.4%;  Score 45;  DB 11;  Length 527;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 99
```

RESULT 150
US-10-301-480-622505
; Sequence 622505, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 622505
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-622505

Query Match 1.4%; Score 45; DB 12; Length 527;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2895 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 2939
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACA 99
```

RESULT 151
US-09-925-065A-765896
; Sequence 765896, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765896
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-765896

Query Match 1.4%; Score 45; DB 7; Length 545;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        291 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 335
```

RESULT 152
US-09-925-065A-799095/c
; Sequence 799095, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 799095
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-799095

Query Match 1.4%; Score 45; DB 7; Length 546;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        210 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 166
```

RESULT 153
US-09-925-065A-615308
; Sequence 615308, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 615308
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-615308

Query Match 1.4%; Score 45; DB 7; Length 549;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 123


RESULT 154
US-09-925-065A-591948/c
; Sequence 591948, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591948
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591948

  Query Match             1.4%;  Score 45;  DB 7;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        157 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 113


RESULT 155
US-09-925-065A-12438/c
; Sequence 12438, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12438
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-12438

  Query Match             1.4%;  Score 45;  DB 7;  Length 563;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 117


RESULT 156
US-10-301-480-113675/c
; Sequence 113675, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113675
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-113675

  Query Match             1.4%;  Score 45;  DB 11;  Length 563;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 117


RESULT 157
US-10-301-480-727084/c
; Sequence 727084, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 727084
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-727084

  Query Match             1.4%;  Score 45;  DB 12;  Length 563;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 117
```

```
RESULT 158
US-09-925-065A-13098
; Sequence 13098, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13098
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-13098

  Query Match             1.4%;  Score 45;  DB 7;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        350 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 394


RESULT 159
US-09-925-065A-338452
; Sequence 338452, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 338452
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-338452

  Query Match             1.4%;  Score 45;  DB 7;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 477


RESULT 160
US-10-301-480-114335
; Sequence 114335, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114335
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-114335

  Query Match             1.4%;  Score 45;  DB 11;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        350 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 394


RESULT 161
US-10-301-480-411598
; Sequence 411598, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 411598
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-411598

  Query Match             1.4%;  Score 45;  DB 12;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 30;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 477


RESULT 162
US-10-301-480-727744
; Sequence 727744, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 727744
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-727744

Query Match 1.4%; Score 45; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||
Db 350 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 394

RESULT 163
US-10-301-480-1025007
; Sequence 1025007, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1025007
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1025007

Query Match 1.4%; Score 45; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||
Db 433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 477

RESULT 164
US-09-925-065A-371064/c
; Sequence 371064, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 371064
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-371064

Query Match 1.4%; Score 45; DB 7; Length 590;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||
Db 433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 389

RESULT 165
US-09-925-065A-927764
; Sequence 927764, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 927764
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-927764

Query Match 1.4%; Score 45; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105

RESULT 166
US-09-925-065A-927765
; Sequence 927765, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 927765
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-927765

  Query Match             1.4%;  Score 45;  DB 7;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105


RESULT 167
US-09-925-065A-949474
; Sequence 949474, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 949474
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-949474

  Query Match             1.4%;  Score 45;  DB 7;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105


RESULT 168
US-09-925-065A-695847/c
; Sequence 695847, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 695847
;   LENGTH: 596
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-695847

  Query Match             1.4%;  Score 45;  DB 7;  Length 596;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2906 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 2950
             |||||||||||||||||||||||||||||||||||||||||||||
Db        508 TGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATAGCGAAACCC 464


RESULT 169
US-09-925-065A-10892/c
; Sequence 10892, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10892
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-10892

  Query Match             1.4%;  Score 45;  DB 7;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 260


RESULT 170
US-10-301-480-112129/c
; Sequence 112129, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112129
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-112129

  Query Match             1.4%;  Score 45;  DB 11;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 260


RESULT 171
US-10-301-480-725538/c
; Sequence 725538, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 725538
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-725538

  Query Match             1.4%;  Score 45;  DB 12;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
             |||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 260


RESULT 172
US-10-301-480-253035/c
; Sequence 253035, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 253035
;   LENGTH: 603
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-253035

  Query Match             1.4%;  Score 45;  DB 12;  Length 603;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         98 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 54


RESULT 173
US-10-301-480-866444/c
; Sequence 866444, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 866444
;   LENGTH: 603
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-866444

  Query Match             1.4%;  Score 45;  DB 12;  Length 603;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db         98 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 54


RESULT 174
US-09-925-065A-160227/c
; Sequence 160227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 160227
```

```
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-160227

  Query Match             1.4%;  Score 45;  DB 7;  Length 604;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         99 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 55


RESULT 175
US-09-925-065A-799096/c
; Sequence 799096, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 799096
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-799096

  Query Match             1.4%;  Score 45;  DB 7;  Length 604;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 202


RESULT 176
US-09-925-065A-795229/c
; Sequence 795229, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 795229
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-795229

  Query Match             1.4%;  Score 45;  DB 7;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         71 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 27


RESULT 177
US-09-925-065A-850785/c
; Sequence 850785, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 850785
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-850785

  Query Match             1.4%;  Score 45;  DB 7;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db         71 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 27


RESULT 178
US-10-301-480-441075/c
; Sequence 441075, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 441075
```

```
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-441075

  Query Match             1.4%;  Score 45;  DB 12;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 389


RESULT 179
US-10-301-480-1054484/c
; Sequence 1054484, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1054484
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1054484

  Query Match             1.4%;  Score 45;  DB 12;  Length 606;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 389


RESULT 180
US-09-925-065A-846491/c
; Sequence 846491, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 846491
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-846491

  Query Match             1.4%;  Score 45;  DB 7;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        399 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 355


RESULT 181
US-09-925-065A-839499/c
; Sequence 839499, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 839499
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-839499

  Query Match             1.4%;  Score 45;  DB 7;  Length 627;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2896 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 2940
              |||||||||||||||||||||||||||||||||||||||||||||
Db        402 GTGGATCACCTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACAT 358


RESULT 182
US-09-925-065A-675355/c
; Sequence 675355, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 675355
```

```
;   LENGTH: 662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-675355

  Query Match             1.4%;  Score 45;  DB 7;  Length 662;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        260 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 216


RESULT 183
US-09-925-065A-675356/c
; Sequence 675356, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 675356
;   LENGTH: 662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-675356

  Query Match             1.4%;  Score 45;  DB 7;  Length 662;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        260 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 216


RESULT 184
US-10-301-480-34229
; Sequence 34229, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34229
;   LENGTH: 713
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-34229

  Query Match             1.4%;  Score 45;  DB 11;  Length 713;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3077 TGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||
Db        468 TGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 512


RESULT 185
US-10-301-480-647638
; Sequence 647638, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 647638
;   LENGTH: 713
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-647638

  Query Match             1.4%;  Score 45;  DB 12;  Length 713;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3077 TGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 3121
             |||||||||||||||||||||||||||||||||||||||||||||
Db        468 TGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCT 512


RESULT 186
US-10-301-480-593973/c
; Sequence 593973, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 593973
;   LENGTH: 723
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-593973

  Query Match             1.4%;  Score 45;  DB 12;  Length 723;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        113 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 69
```

RESULT 187
US-10-301-480-1207382/c
; Sequence 1207382, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1207382
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1207382

Query Match 1.4%; Score 45; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 113 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 69

RESULT 188
US-10-301-480-572562
; Sequence 572562, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 572562
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-572562

Query Match 1.4%; Score 45; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 560 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 604

RESULT 189
US-10-301-480-1185971
; Sequence 1185971, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1185971
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1185971

Query Match 1.4%; Score 45; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 560 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 604

RESULT 190
US-09-925-065A-13524/c
; Sequence 13524, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13524
; LENGTH: 793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-13524

Query Match 1.4%; Score 45; DB 7; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696

RESULT 191
US-09-925-065A-77588/c
; Sequence 77588, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77588
;   LENGTH: 793
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-77588

  Query Match             1.4%;  Score 45;  DB 7;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696


RESULT 192
US-10-301-480-114761/c
; Sequence 114761, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114761
;   LENGTH: 793
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-114761

  Query Match             1.4%;  Score 45;  DB 11;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696


RESULT 193
US-10-301-480-178827/c
; Sequence 178827, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 178827
;   LENGTH: 793
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-178827

  Query Match             1.4%;  Score 45;  DB 11;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696


RESULT 194
US-10-301-480-728170/c
; Sequence 728170, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 728170
;   LENGTH: 793
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-728170

  Query Match             1.4%;  Score 45;  DB 12;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
             |||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696


RESULT 195
US-10-301-480-792236/c
; Sequence 792236, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 792236
;   LENGTH: 793
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-792236
```

```
  Query Match             1.4%;  Score 45;  DB 12;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        740 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 696


RESULT 196
US-10-301-480-556511/c
; Sequence 556511, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 556511
;   LENGTH: 803
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-556511

  Query Match             1.4%;  Score 45;  DB 12;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        186 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 142


RESULT 197
US-10-301-480-1169920/c
; Sequence 1169920, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1169920
;   LENGTH: 803
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1169920

  Query Match             1.4%;  Score 45;  DB 12;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        186 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 142


RESULT 198
US-10-301-480-556318/c
; Sequence 556318, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 556318
;   LENGTH: 818
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-556318

  Query Match             1.4%;  Score 45;  DB 12;  Length 818;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        186 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 142


RESULT 199
US-10-301-480-1169727/c
; Sequence 1169727, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1169727
;   LENGTH: 818
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1169727

  Query Match             1.4%;  Score 45;  DB 12;  Length 818;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        186 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 142


RESULT 200
US-09-925-065A-88887
; Sequence 88887, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
```

; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 88887
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-88887

Query Match 1.4%; Score 45; DB 7; Length 896;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
              |||||||||||||||||||||||||||||||||||||||||||||
Db        457 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 501
```

Search completed: May 11, 2006, 06:25:33
Job time : 1350 secs

THIS PAGE BLANK (USPTO)